



TITLE:

環境データの組織化と環境問題への適用に関する研究(Dissertation\_全文)

AUTHOR(S):

藤原, 正弘

---

CITATION:

藤原, 正弘. 環境データの組織化と環境問題への適用に関する研究. 京都大学, 1986, 工学博士

ISSUE DATE:

1986-03-24

URL:

<https://doi.org/10.14989/doctor.r5884>

RIGHT:

環境データの組織化と環境  
問題への適用に関する研究

昭和 60 年 9 月

藤 原 正 弘

# 環境データの組織化と環境 問題への適用に関する研究

昭和 60 年 9 月

藤 原 正 弘

# 環境データの組織化と環境問題への適用に関する研究

## 目 次

緒 論 .....	1
第 1 節 環境データの特徴とそれに起因する問題点 .....	1
第 2 節 本論文でとり上げる問題 .....	2
第 I 編 環境データの組織化に関する考察 .....	5
第 1 章 大気環境データの組織化 .....	7
第 1 節 概 説 .....	7
第 2 節 データ組織化の意義 .....	7
第 3 節 大気環境データ組織化の課題 .....	9
3.1 課題の抽出 .....	9
3.2 データの処理、蓄積方式についての課題 .....	10
第 4 節 データの蓄積方式の標準化 .....	15
第 5 節 結 語 .....	25
第 2 章 水質及び廃棄物データの組織化 .....	26
第 1 節 概 説 .....	26
第 2 節 水質環境データの組織化 .....	26
2.1 データの測定及び処理の現状 .....	26
2.2 データの組織化のための課題 .....	29
2.3 データの蓄積方式の標準化 .....	30
第 3 節 廃棄物データの組織化 .....	37
3.1 データ観測の現状 .....	37
3.2 データ管理の方向 .....	40
第 4 節 結 語 .....	44



第3章 国内外における環境情報システムの現状 .....	46
第1節 概 説 .....	46
第2節 わが国における環境情報システムの現状 .....	46
第3節 諸外国における環境情報システムの現状 .....	50
3.1 米 国 .....	50
3.2 カ ナ ダ .....	58
3.3 西 独 .....	59
3.4 スウェーデン .....	60
第4節 結 語 .....	61
第4章 環境データベースの作成と環境問題への適用 .....	62
第1節 概 説 .....	62
第2節 環境データベースの意義 .....	62
2.1 データベースの定義 .....	62
2.2 環境施策とデータベースの関係 .....	63
2.3 環境データベースの位置づけ .....	64
第3節 環境データベースの作成 .....	66
3.1 作成するデータベースの機能 .....	66
3.2 データファイルの作成 .....	69
第4節 環境データベースの利用 .....	73
4.1 パラメータ・ユーザによる利用例 .....	73
4.2 プログラミング・ユーザによる利用例 .....	75
第5節 データベースの環境問題への適用上の課題 .....	77
第6節 結 語 .....	78
第Ⅱ編 環境データの環境問題への適用技術に関する研究 .....	79
第5章 データの自動チェックシステムに関する研究 .....	81

第1節	概 説	81
第2節	環境データの自動チェックの課題	81
2.1	現行システムの問題点	81
2.2	自動チェックの環境データへの適用の方向	83
第3節	データチェック手法の有効性の検討	84
3.1	データチェック手法の種類及び検討対象	84
3.2	有効性の評価尺度	86
3.3	ADFの適用の検討	87
3.4	ADF法と加重平均法、偏差法との比較	92
3.5	項目間回帰法、上下限值法の有効性の検討	95
第4節	結 語	98
第6章	データの2次元化手法に関する研究	100
第1節	概 説	100
第2節	スプライン法	100
2.1	特 徴	100
2.2	計算方法	101
2.3	計算上の条件	103
2.4	パラメータ $\sigma$ の決定方法	104
第3節	補間の適用性の検証方法	105
3.1	近似分布の決定	105
3.2	近似分布推定法の妥当性	108
第4節	実データへの適用	109
4.1	パラメータ $\sigma$ の決定	110
4.2	有効性の検討	110
第5節	結 語	114
第7章	大気汚染測定局の最適配置に関する研究	116

第1節	概 説	116
第2節	測定局の最適削減法	117
第3節	削減局決定のための方法	118
第4節	実データへの適用	120
第5節	結果の検討	122
第6節	結 語	127
第8章	水質環境データの2次元化と測定点の合理的配置に関する研究	128
第1節	概 説	128
第2節	スプライン法の水質汚濁データへの適用	129
2.1	スプライン法と水質汚濁データへの適用上の課題	129
2.2	対象領域、境界条件、初期値の設定	129
2.3	収束条件の検討	130
2.4	収束性の改良	133
2.5	パラメータ $\sigma$ の値の検討	134
第3節	スプライン法による琵琶湖南湖の水質濃度分布パターンの推定	136
第4節	合理的測定点配置の検討	137
4.1	考え方	137
4.2	測定点の分布曲面への影響度	138
4.3	測定点削減による分布曲面の変化	142
4.4	削減可能測定点の選定及び代表測定点の選定	146
第5節	結 語	149
結 論		152
使用記号		156
参考文献		160

## 緒 論

### 第 1 節 環境データの特徴とそれに起因する問題点

環境問題に関する各種データ（以下、環境データという）は、環境科学研究や環境施策の実施のために不可欠なものである。大気汚染や水質汚濁のデータは、モニタリングシステムの整備が全国的に進んだことにより質的量的に相当なレベルで把握されている。このようなデータを環境科学研究や環境施策の実施に有効に利用するようにするためには、総合的にデータを管理するシステムが必要である。最近、データベース関連の技術はめざましいものがあり、環境データの総合的管理は技術的に特別むずかしいものとは考えられない。しかし、他のデータとは異なるその特徴には十分留意しなければならない。環境データの特徴を、例えば大気汚染や水質汚濁などの環境汚染の測定データを対象に考えると次のような点があげられる。

- ① データは時間によって変わる（物性データのように時間的に不変なものではない）。
- ② データは地域（Area）の汚染を知るために測定されるのであるが、個々のデータはある地点（Point）で測定されたものである。
- ③ 地域または水域の汚染度測定のデータは測定点の配置に依存する。
- ④ データには測定機関従属性（データの内容が測定者に依存し、測定者のみが知っている属性情報があるなどの特性）がある。
- ⑤ データは誤差を有し、かつこれが不可避である。

このような環境データの特徴に起因して生じるデータ管理、利用上の問題点を整理すると次のようである。

まず、第 1 点はデータの組織化の問題である。すなわち、環境汚染（例えば、大気汚染、水質汚濁など）は行政区域をこえて広域的であることが多いが、その監視測定は各行政区域ごとに関係機関が実施している。そしてとられたデータは当該機関ごとにそれぞれの方式で記録され保有されている。広域的な汚染問題の解析には、それらを集めて総合的に行うことが必要である。別々に測定され管理



されているデータを同一の解析方法の中にとり込むためにはデータ処理に関する少なからぬ作業を必要とする。環境研究（または環境施策）のプロジェクトごとにこの作業を行う従来の方式から、個々に測定されたデータを中央の機関に集中し、統一的な方法で処理、管理し、ユーザが利用しやすいようにする方式への変換が望まれる。

第2点目は、データ処理技術上の問題である。この代表的な課題としてデータの誤差の処理の問題がある。環境データの誤差は測定時に発生するもの、伝送時に発生するもの、その他の人為的ミスに基づくものなどである。測定技術の精度には限界があるので、これによる一定の誤差が生じることはやむを得ない。各種の原因によって発生する異常値の混入は当該データが所属するデータ群の信頼性及び精度を落とすので、適当な方法でチェックすることが必要である。

第3番目は地域の汚染度の表現に関する問題である。すなわち、環境データは地域の汚染度を知るためにとられるものであるが、個々のデータはそれぞれの地点で測定されたもの（点データという）である。一方、環境問題の解析対象となる他のデータ（例えば人口データ、自然度データなど）はある地域についてのデータ（面データという）であるので、環境データも面データとして表現されることがデータ解析上必要なことが生じうる。このことは一歩進んで、測定点の配置問題に関係してくる。環境データの測定結果は測定点をどこにとるかによって変わってくる。環境測定は汚染度の高い地域を中心に始められ、社会的要請の強まりに応じて周辺地域に広げられて来ている。また、汚染物質に対する社会的関心の変化によって測定点の配置が変えられることもある。このようなことに対応するためには、測定点の配置問題を検討することにより地域の汚染度を表わすデータの意義づけをすることが必要である。

## 第2節 本論文でとり上げる問題

前述のような環境データが有するいくつかの問題点についてその対応策を検討するのが本論文の主題である。まず、環境データの組織化——以下において、この言葉はコンピュータによるデータの総合的管理の実施により環境データの有効

な利用を可能ならしめるためにデータの観測、移送、評価、編集、蓄積、管理の各段階について、データが有する問題点を解決し、全体として総合化され、利用しやすい形のデータを作成することを意味するものとする。——について検討し、ついで組織化されたデータをもとにしてデータベースを作成する。その他の問題点（データの誤差の問題、面データへの変換の問題、測定点の配置に関する問題）には技術的手法にもとづく対応策を提案する。すなわち本論文の検討対象は次のようである。

- ① 環境データの組織化のための課題の抽出と対応策の検討
  - ② 環境データベース（環境データを対象とするデータベース、以下同じ）の作成と環境問題への適用にあたっての課題の検討
  - ③ データの環境問題への適用のために必要とされる技術的手法（④～⑥）の検討及びその実際問題への適用
  - ④ データの自動チェックシステム
  - ⑤ スプライン補間法を用いたデータの２次元化手法
  - ⑥ 既存の環境汚染測定網を対象にした合理的測定点配置の検討手法
- 上記①～②の内容は第Ⅰ編（第１～４章）で、③の内容は第Ⅱ編（第５～８章）で論じる内容である。

以下、章別に検討内容を簡単に述べる。

まず、第１章では、大気環境データについて組織化の課題の指摘と対応策の検討を行うとともに磁気テープへのデータ記録方式の標準化については具体的な対応策を提案する。第２章では、同様の内容を水質環境データについて行うとともに廃棄物データについてはそれに関する情報管理の現状と今後の方向の考察を行う。ついで、第３～４章では環境データベースについて検討を行う。まず第３章では、わが国および先進諸外国における環境データベースの開発、運営の状況をレビューする。第４章では、第１～２章の検討結果にもとづき実際にデータベースを作成し、これを用いることにより、データの処理利用がどのように便利にかつ効率的になるかを示す。また、データを実際の環境問題に適用するうえで必要な技術的手法を指摘する。

次に、第Ⅱ編では、データの環境問題への適用のために必要とされる技術手法のいくつか——データチェック手法、データのメッシュデータ化手法、測定点の最適配置の検討手法——について検討する。そして実際の環境問題に適用することにより、その実用性を検討する。すなわち、第5章では、大気環境データを用いてコンピュータによるデータチェックの自動化の各種方法を比較検討し、適当なデータチェック手法を提案する。第6章では、データの2次元化のためのデータ補間法としてのスプライン法について述べる。ここではデータ補間法の環境汚染濃度への補間の適用性の検証方法を提案し、これを実際のデータに適用することにより、スプライン法が大気汚染濃度の補間によく適用しうるものであることを証明する。第7章では、測定局の最適配置の検討のための一つの手法として、スプライン法によるデータ補間法を用いることにより既存モニタリングシステムの再検討方法——必要最小測定局数の決定と測定局の削減方法——を論じる。第8章ではスプライン補間法を水質環境データに適用し、前二章と同様の検討を行う。測定点の合理的配置については、各測定点で複数の項目を測定対象としている場合について検討し、これを既存モニタリングシステムの再検討に適用する。

以上のように、本研究は、環境データの特性に起因する問題に留意し、その技術的対応策を検討するものであり、検討結果は環境データの作成、データチェックシステム、データの2次元化方法、測定点の最適配置方法の検討の際に技術的方策を与えるものとなろう。

## 第Ⅰ編 環境データの組織化に関する考察

# 第 1 章 大気環境データの組織化

## 第 1 節 概 説

環境データは全国各地で多くの機関で測定されており、またその量は膨大なものである。全国各地で測定されたデータを集めて、総合的に利用しようとする場合に必要な条件は、個々のデータが正しく測定され、処理され、そしてその全体が相互に有機的な関連を有する集合体としてのデータとなっていることである。これはデータの組織化という言葉でいいかえることができる。本章では、まず、環境データを組織化することの意義について論じ、ついで、大気環境データ（一般環境の大気汚染測定データをいう、以下同じ）の測定、チェック、編集、蓄積、管理に関してどのような問題があるかを分析する。そしてデータの組織化の重要な一つの課題であるデータの記録方式の標準化について具体的方策を提案する。

## 第 2 節 データの組織化の意義

データの組織化は具体的にどのような作業を通じて実現するか、またその意義は何かを見てみる。

図 1 に見られるように、あるデータ（データ\*）がある機関（機関  $\alpha$ ）におけるデータ組織化の作業を通じてデータ\*\*に加工されたというフローを考える。データ\*は各測定者、測定機関が測定したデータを機関  $\alpha$  に提供したデータで、まだ十分組織化されていないデータである。データ\*\*はデータ\*を処理、加工することにより組織化されたデータに変わったものである。機関  $\alpha$  はここではデータセンターのような機関を考えればよいが、地方自治体が測定したデータを中央行政機関が制度上収集するような場合、その中央行政機関を考えればよい。データ\*\*が十分組織化されたデータとなっていることを前提にするとデータ\*とデータ\*\*がどのように相違するかは表 1 のように表わすことができる。すなわち、データ\*\*はデータの全体を総合的に解析するに必要な情報規則を正しく有しているデータであるといえることができる。この組織化の作業は機関  $\alpha$  がすべて行うとは限らない。各測定者（又は測定機関）が予め定められた一定の共通の方法で

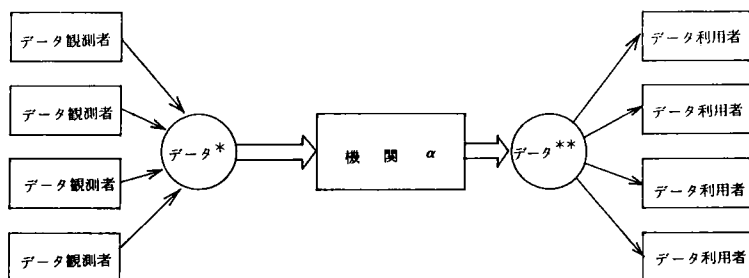


図 1 未組織化データの既組織化データへの変換フロー

表 1 未組織化データと既組織化データの相違

		デ ー タ *	デ ー タ **
属 性 情 報	(例) WHO (測定者) WHERE (測定場所) WHEN (測定時間) WHAT (測定項目) HOW (測定方法)	<ul style="list-style-type: none"> <li>属性情報が欠落している。 (例) 測定場所に関する情報がない。</li> <li>属性情報の表示方法が不統一 (例) 測定場所の情報の表示が市町村名表示、メッシュコード表示、座標軸表示であったりする。</li> </ul>	<ul style="list-style-type: none"> <li>欠落したデータは付加されている。</li> <li>不統一な表示方法は統一的な表示方法に変更されている。</li> </ul>
データ内容	(例) データチェック法 異常値のとりあつかい方 測定機の保守管理方法 サンプリング方法	<ul style="list-style-type: none"> <li>データチェックがなされていない。</li> <li>データチェックが統一的な方法で行なわれていない。</li> <li>異常値が混入している。</li> </ul>	<ul style="list-style-type: none"> <li>データチェックが統一的な方法でやりなおされている</li> <li>異常値は混入していない。</li> </ul>
データ形態	(例) 有効ケタ数 単 位 項目コード、単位コード、 データ配列、フォーマット	<ul style="list-style-type: none"> <li>有効ケタ数、単位、コードは統一されていない。</li> <li>データの配列順序及びフォーマットがまちまちである。</li> </ul>	<ul style="list-style-type: none"> <li>有効ケタ数、単位、コードは統一的なものになっている。</li> <li>データの配列順序及びフォーマットは統一したものにつくりなおされている。</li> </ul>



データ処理、加工、整理、収録を行なった上、機関 $\alpha$ に提出（又は提供）することとも考えられる。一定の共通な方法の設定はデータの組織化の1つの重要な課題である。国家的な規模のデータベースの作成もこのようなデータを前提として可能である。また、このように共通な方法でデータを処理することにより、データ解析のプログラムが共通に利用できるので、個々の測定者（又は測定機関）がそれぞれこれを作成する必要がなくなり、各測定者（又は測定機関）にとっても有益である。

### 第3節 大気環境データ組織化の課題

#### 3.1 課題の抽出

データの観測から処理、蓄積そして利用に至るデータの流れの各段階ごとにまずデータの組織化の課題<sup>1)</sup>を抽出する。

- ① データ観測段階：各自治体の大気環境データの測定は法制度上の規定<sup>2)</sup>に基づいて行われており、かつ大部分が機械化されているので、全国的に比較的標準化されていると言える。しかし、測定機の保守管理の方法は十分標準化されているとはいいがたい。これについては環境庁の検討<sup>3)</sup>のほかに、東京都などの自治体における検討<sup>4)</sup>も進められている。データの地域代表性（測定局のデータが地域の汚染の状態をよく代表したものであるかどうかについての問題）は、測定局の最適配置という観点から検討が必要である。
- ② データの移送段階：各自治体でとられたデータを他に移送したり、中央に集めたりする場合、磁気テープを介して行われることが多い。しかし、データの磁気テープへの記録方式が自治体によって異なっているので、当該自治体以外の機関がそのデータを利用することは容易でない。そこで、各自治体における磁気テープへのデータの記録方式（磁気テープの仕様）を標準化することが課題である。
- ③ データチェック段階：自動測定機で測定されたデータは適当な方法でチェックし、異常値の判定、修正が行われねばならない。疑わしいデータが観測された場合に、これを異常値とするかどうかの判断方法は一般化されていない。人手

による異常値の判定、修正等の作業は担当者の経験にもとづく判断にゆだねられている場合が多く<sup>5)</sup>、チェックされたデータの精度に差が生じるものとなる。

- ④ データの編集・入力段階：データの配列、有効桁数、属性情報（測定位置、方法等に関する情報）等を整理、統一すること、及び使用頻度の高い統計量（平均値、最大値、98%値など）を集計、編集することが必要である。
- ⑤ データの蓄積・管理の段階：多量のデータを標準化された方式で蓄積し、コンピュータシステムの支援によりユーザの希望に応じて随時提供が可能のように管理することが必要である<sup>6)7)</sup>。

### 3.2 データの処理、蓄積方式についての課題

全国の自治体における大気環境データシステムの中でデータ処理、蓄積の実態はどのようなかを調べた。

筆者らが行った調査<sup>8)</sup>（調査は、昭和51年4月時点の状況を、47都道府県と大気汚染防止法による60政令市の合計107自治体を対象にアンケート方式で実施した。アンケートの回収率は100%であった。）によると、次のようである<sup>9)</sup>。

表2に示すように、大気環境データは、合計55自治体において、磁気テープ（MT）に収録され保有されている。それ以外の自治体の大部分では、1月単位のデータ表（月報リスト）の形でデータを保有している。データチェック及びデータ修正の作業（以下データチェックと呼ぶ。）は、その方法に差はあれ、大部分の自治体で実施されている（表3）。自治体のコンピュータには、その機能上、モニタリングシステム用のもの（以下モニタリング用コンピュータと呼ぶ。）とデータ処理やデータ解析に用いるもの（以下データ処理用コンピュータと呼ぶ。）とに分類できる。後者についての調査結果は次のとおりである。コンピュータの規模は、内部記憶容量が512キロバイト以下または256キロキャラクタ以下という非常に小型のものがほとんどである（表4）。用いられているコンピュータは種類が多様であり、したがって、それに依存する磁気テープのトラック数、記録密度は多様である。そのうちで9トラック、800BPIのものが半数をしめ

表2 大気環境データの記録媒体別自治体数

自治体区分 記録媒体	MTを保有し ている自治体 (55)について	MTを保有して いない自治体 (52)について	計 (107)
1. 測定チャート	22	42	64
2. 月報リスト	50	50	100
3. 紙テープ	7	10	17
4. 磁気テープ	55	0	55
5. マイクロフィルム	1	0	1
6. マイクロフィッシュ	1	0	1
7. その他	3	4	7

表3 データチェックの実施有無別自治体数

自治体区分 記録媒体	MTを保有し ている自治体 (55)について	MTを保有して いない自治体 (52)について	計 (107)
1. データチェックして いる	54	48	102
2. データチェックして いない	1	4	5

表4 コンピュータの内部記憶各量的自治体数

単位 ランク	自治体数(50)		
	Kバイト	Kワード	Kキャラクタ
1～16	2	3	0
17～32	3	12	3
33～64	5	1	1
65～128	1	0	2
129～256	7	2	2
257～512	3	0	1
513～	2	0	0
計	23	18	9

注) キロバイト, キロワード, キロキャラクタは, 内部記憶容量の単位であり, 概略的には, 1キャラクタ=0.75バイト, 1ワード=2バイト(32キロワード以下の場合), 1ワード=4バイト(32キロワードをこえる場合)である。

ており（表５）。また、７トラック、５５６ＢＰＩというような小型コンピュータ特有のものも少なくない。記録コードも、用いられているコンピュータの種類が多様であることから、EBCDIK, BCD, ISO, BINARY, ASCII など多くの記録コードが用いられており、EBCDIKが比較的多いがとくに多く使用されている記録コードを特定することはできない（表６）。磁気テープにおけるレコードサイズは、１００バイト以下の短いレコードサイズの自治体が１２件あるが、２０００バイト程度のレコードサイズのものまで各自治体によって種々である（表７）。データ配列は、おおまかなパターン分類をすると表８に示す１７通りのデータ配列になるが、やはり、自治体によって種々である<sup>9)</sup>。表９は、代表的な自治体の磁気テープ仕様の概要<sup>10)</sup>を示したものである。レコードサイズ、データ配列、記録コード等はそれぞれ各自治体によって異なり、共通性が余りないことが知られる。

以上から、わが国の大気環境データの処理についての課題として次の点が指摘される<sup>9)~11)</sup>。

- ① テレメータシステムによる大気汚染モニタリングシステムを通して、データを磁気テープに記録している自治体が多い。しかし、用いられているコンピュータは、モニタリング用コンピュータにしる、データ処理用コンピュータにしる、比較的小型のものである。この点が地方自治体における大気環境データ処理のひとつの注意すべき特徴である。
- ② データのチェックは、ほとんどすべての自治体で行なわれているが、人手によるチェックであり、その方法は統一された方法ではない。そのため、全国的に見た場合の質の均質性については問題がある。
- ③ 各自治体では各種のコンピュータが用いられているので、ハードウェア（コンピュータ）に依存する磁気テープ仕様（トラック数、記録密度、記録コードなど）が自治体ごとに異なる。この統一はハードウェアの仕様に影響するものであり困難である。
- ④ ハードウェアに依存しない磁気テープ仕様（レコードサイズ、データ配列等のフォーマットなど）についても自治体ごとに異なっており、その種類は事実

表5 磁気テープのトラック数、記録密度別自治体数

トラック数	記録密度	自治体数
9トラック	800 BPI	26
	1600 BPI	13
7トラック	800 BPI	9
	556 BPI	3
計		51

注) BPIとは、記録密度の単位で、  
バイト/インチ  
(byte per inch) のことである。

表6 コンピュータの記録コード別自治体数

コード	自治体数
EBCDIK	19
BCD	7
ISO	6
BINARY	6
ASCII	5
その他	10
計	53

表7 レコードサイズ別自治体数

レコードサイズ	自治体数	
	ワード	バイト
ワード (または ~ 100バイト)	4	12
100 ~ 300	4	10
300 ~ 500	1	3
500 ~ 1000	6	2
1000 ~ 2000	2	0
2000 ~	0	2
不明	7	
	53	

表8 データ記録順序別自治体数

種類	データの配列・順序	自治体数
1	月>日>時>局>▼項目	8
2	月>日>時>▼局>項目	7
3	局>月>日>時>▼項目	6
4	局>月>日>▼時>項目	2
5	局>月>日>時>▼項目	4
6	月>日>局>▼時>項目	1
7	局>項目>月>▼日>時	3
8	項目>局>月>日>▼時	5
9	月>局>項目>▼日>時	2
10	月>局>日>時>▼項目	1
11	月>日>局>▼項目>時	1
12	月>日>局>項目>▼時	3
13	月>日>時>▼項目>局	2
14	局>項目>▼月>日>時	1
15	月>日>項目>局>▼時	1
16	局>月>日>項目>▼時	1
17	月>日>時>局>項目▼	2
18	その他、記入無し	3
	合計	53

(>記号: 例 (月>日) とは1月1日のデータ, 1月2日のデータ, 1月3日のデータ……の順)  
(▼はレコードの切れ目を示す)

表9 主要自治体における磁気テープ仕様の概要(昭和51年度)

自治体	1論理レコードに収録されたデータ	レコードサイズ、 ブロッキング	記録コード	年間データの ファイル数	電算機種	その他
埼玉県	(1局, 1項目, 24時間) 1論理レコード =1物理レコード	80バイト/レコード 40レコード/ブロック (ただしブロック長 3204バイト)	EDCDIK	1	ACOS (TOSBAC) 600	
千葉県	(1局, 1項目, 1カ月) 1論理レコード= =4物理レコード	512バイト/レコード ブロッキングなし	2進 (符号なしパック)	12	TOSBAC 40C	非標準ラベルあり
東京都	(1局, 1項目, 24時間) 1論理レコード =1物理レコード	107バイト/レコード 50レコード/ブロック	EBCDIC (ゾーン付)	1	NEAC 2200-400	他のフォーマットも所有する
神奈川県	(1局, 全項目, 1時間) 1論理レコード =1物理レコード	1215バイト/レコード ブロッキングなし	EBCDIK (パック形式と パックなしの 並用)	12	NEAC M-4	
愛知県	(1局, 1項目, 1カ月) 1論理レコード =1物理レコード PIマスタ(汚染質), WIマスタ(気象)の2 ファイル	2260バイト/レコード (PIマスタ) 2330バイト/レコード (WIマスタ) 1レコード1ブロック	EBCDIK	2	HITAC 8400	
大阪府	(1局, 1項目, 1カ月) 1論理レコード =7物理レコード	505バイト/レコード 516バイト/ブロック	2進(ただし データは4バ イト構成)	1	ACOS (NEAC) 500	
大阪市	(1局, 全項目, 1時間) 1論理レコード =1物理レコード	100バイト/レコード 10レコード/ブロック	EBCDIK (ゾーンとパッ クの共用)	2	HITAC 8250	非標準のラベルあり
兵庫県	(1局, 1項目, 1カ月) 1論理レコード =1物理レコード	801ワード/レコード 1ワードは16ビット (2バイト)ブロッキ ングなし	データ部は2 進, キー部は ASCII	1	HIDIC 700	ラベルなし
神戸市	(1局, 1項目, 24時間) 1論理レコード =1物理レコード	128バイト/レコード 20レコード/ブロック	EBCDIK	1	FACOM 230	
岡山県	(1局, 1項目, 24時間) 1論理レコード =1物理レコード	100バイト/レコード 10レコード/ブロック	EBCDIK	1	NEAC M-4	ラベルなし
石川県	(1局, 全項目, 1時間) 1論理レコード =1物理レコード	48バイト/レコード 2304バイト/ブロック	パケットデシ アルと2進の 並用	12	FACOM U-200	ラベルなし
福井県	(1局, 全項目, 1時間) 1論理レコード =1物理レコード	60バイト/レコード 1440バイト/ブロック	2進	12	JAC 120	
大分県	(全局, 全項目, 1時間) 1論理レコード =1物理レコード	512バイト/レコード 512バイト/ブロック	EBCDIK とパ ケットデシ アルの並用	12	FACOM U-400	
鹿児島県	(全局, 全項目, 1時間) 1論理レコード =1物理レコード	900バイト/レコード 900バイト/ブロック	ISO	1	OKITAC 4300C	ラベルなし



上自治体の数だけある。しかし、これはハードウェアに依存しないので統一を図ることは比較的容易である。

#### 第4節 データの蓄積方式の標準化

前述のように、大気環境データの処理の上で、各自治体の磁気テープ仕様の多様性が問題になっている。以下においてはこの標準化についての具体案を検討する。磁気テープの仕様の事項には

- ① 記録密度、トラック数、記録コードなど
- ② データの配列（レコードフォーマット、レコードサイズなど）
- ③ 各種コード（測定局コード、測定項目コード、単位コード、欠測コードなど）

がある。①はコンピュータに依存する事項であり、②および③はコンピュータに依存しない事項である。①の事項を標準化することは、コンピュータの仕様の標準化、統一化を意味するので困難である。しかし、この事項（記録密度、トラック数、記録コード）は用いられている種類がそれぞれ数種類に限られる（表5、表6）ので、予め想定されるケースについてコンバージョンプログラムを作成しておき、これにより標準的なものに変換すればよい。一方、②、③については、その種類が無数であり、全てのケースに適用しうるコンバージョンプログラムを作成することは不可能である。したがって、データ作成者の方で予め指定された仕様のものを作成することが望ましい。そこで、以下において②と③の事項についての標準化を検討する。標準仕様の磁気テープが有すべき必要条件として次の事項があげられる。

- ㊦ 記憶容量が小さい小型コンピュータから大型コンピュータまでできるだけ広範囲の機種で使用可能なこと。
- ㊧ モニタリングシステムの変更（測定局数、測定項目数の変化など）があってもフォーマットに変更を生じないものであること
- ㊨ 多量のデータが収録できること
- ㊩ データ処理速度が速いこと

㊦ 各自治体において標準仕様への変換が容易なこと

㊧ ユーザにとって理解しやすいものであること。

これらの必要条件に留意しレコードフォーマット、レコードサイズ、測定局コードなどについて検討を加える。

#### (1) レコードフォーマットの検討

データファイルの中の測定データが有すべき情報は、それ自身のほかに 4 W 1 H (WHAT, WHEN, WHERE, WHO, HOW)に関する属性情報であるが、測定データが属性情報を常にかかえているとなると、データフォーマットは概念的には図 2 の上半分に示すような形態になり<sup>12) 13)</sup>、属性情報の部分が大きなものになってしまい、データ 1 つの記憶に非常に大きい記憶容量を必要

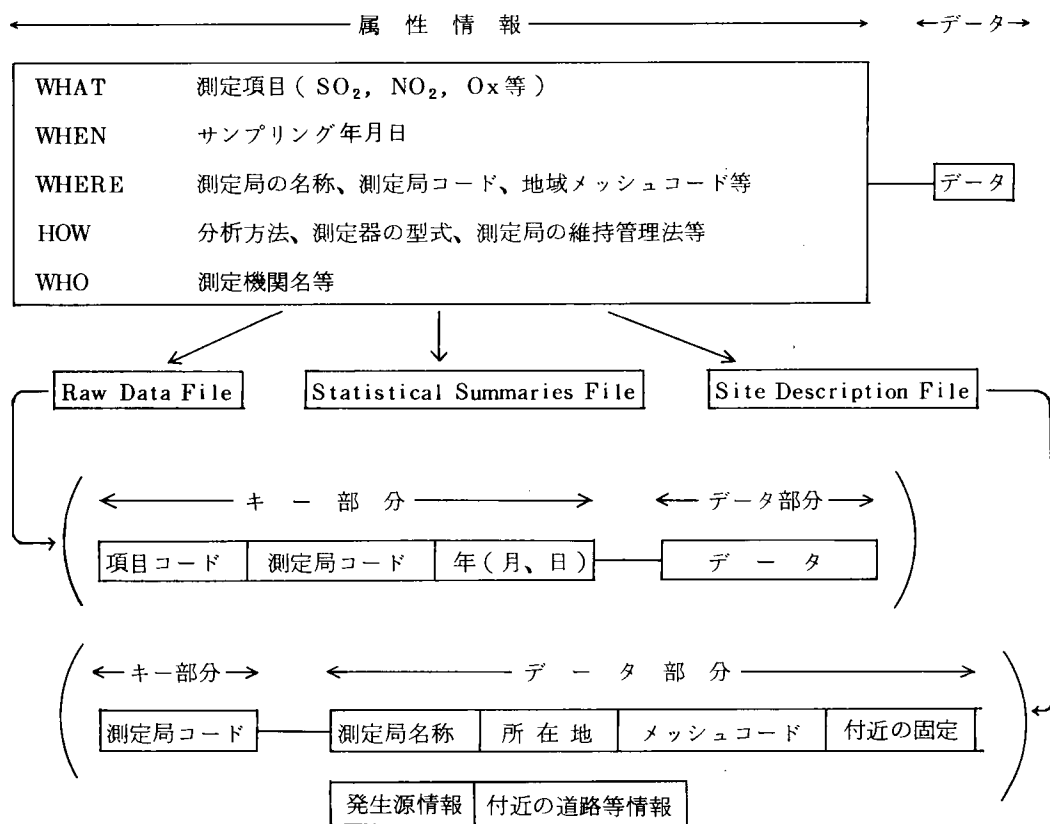


図 2 大気環境データファイルの構成とフォーマットの概念図

とすることとなる。しかし、大気環境データはタイムシリーズになっているので、WHENの情報は個々のデータには付加する必要がない。WHATの情報（測定項目）、WHEREの情報（測定局の場所）、HOWの情報（測定方法）、WHOの情報（測定機関）は、あるデータのグループの中では共通であり、このようなデータグループに相当する大きさのレコードを設定すれば、レコード単位に情報を付加すればよく、データ個々には付加する必要がない<sup>12) 13)</sup>。WHEREの情報は、測定局コード、測定局の名称、所在地、そのメッシュコード、付近の固定発生源の情報、付近の道路等の情報などかなり膨大なものになるので、この情報の付加は測定局コードのみとし、その他は別ファイル(Site Description Fileと呼ぶ。)に入力する。測定データを入力するファイル(Raw Data Fileと呼ぶ。)とこのファイルは測定局コードで連結される。なお、利用上の便宜のため、Summary Fileと呼ばれるもう一つ別ファイル（測定値から計算した統計値をまとめたファイル）を作成しておくことが望ましい。

次に、データの配列の順序について検討する。現在各自治体により採用されているデータ配列順序は表8で示したとおり非常に多様である。しかし、1レコードの中に収録されるデータのタイプで分類してみると、表10に示すように大きく分けて4つのタイプがある<sup>9)</sup>。このうちタイプ3とタイプ4は1レコード中に全測定項目または全測定局のデータが収録されるというものであるので、測定項目や測定局が追加されるとデータ配列を変更せねばならないこととなり、前述の必要条件④に適合しないので採用しがたい。

表 10 1レコード中に記録されるデータ内容の類型

タイプ	測定局	測定項目	測定時間
1	1 局	1 項目	1 日
2	1 局	1 項目	1 月
3	1 局	全項目	1 時間
4	全 局	全項目	1 時間

標準的なレコードフォーマットの概念図を図2の下半分に示した。

## (2) レコードサイズの検討

レコードサイズについては、記憶容量の小さいコンピュータでも利用できるためには、小さい方がよいが、一方、大量のデータをコンパクトに磁気テープに収録できるようにするためにはレコードサイズは大きい方がよい。なお、レコードサイズが大きい方が一般にデータ処理速度が速くなる。これらの点を考慮して最適のレコードサイズが考えられるべきであるが、一つの点を満足させれば他の点が満足しなくなるので、目的に応じたレコードサイズを複数採用せざるを得ない。その1は、中小型コンピュータのことは考慮せず、データの収録性とデータ処理の迅速性について効率を最大にするように考えた場合、その2は自治体がデータ処理用に用いているコンピュータで使用可能であり、かつデータ収録性とデータ処理の迅速性も比較的よいように考えた場合のものである。その3はデータの収録性とデータ処理の迅速性についての不利は許容し、自治体で用いられている比較的小さいコンピュータでも使用が可能なように考えた場合である<sup>12) 13)</sup>。このような前提に基づくと、以下のような3つのレコードサイズの案<sup>12) 13)</sup>が考えられる。

レコードサイズ案1：レコードのデータ部に入れるデータは、1測定局、1測定項目、1年間（4月1日～翌年3月31日）のデータ（ $365 \times 24 = 8760$ データ）である。

レコードサイズ2案：レコードのデータ部に入れるデータは、1測定局、1測定項目、1か月（1日～31日）のデータ（ $31 \times 24 = 744$ データ）である。これは表10中のタイプ2に相当する。

レコードサイズ3案：レコードのデータ部に入れるデータは、1測定局、1測定項目、1日（1時～24時）のデータ（24データ）である。これは表10のタイプ1に相当する。

## (3) コードの設定

コンピュータによるデータ処理を容易にするためには、属性情報をできるだけコードにより表現することが望ましい。コードにより表現することができる

属性情報には、測定局、測定項目、測定データの単位、測定局所在地の用途地域などがある。これらを表現するコードは全国的に統一化されたものが用いられることが望ましいので、これらについての標準コードの設定を検討する。

標準コードの有すべき条件としては、そのコードにより対象（測定局や測定項目など）が唯一のものとして特定できること、更新や廃止などがあってもその部分の変更のみで済み他へ影響を及ぼさないこと、数字（又はアルファベット）であることがあげられる。また有することが望ましい条件として、簡潔であること、使用しやすいこと、コードを見ただけで属性情報の内容がある程度判断できることがあげられる。

#### a) 測定局コード

測定局コードには次のような情報が内蔵されるべきである。

- ① 測定局の所在地
- ② 測定局の管理者
- ③ 測定局の種別
- ④ 高度別の測定

これらの情報のうち①については都道府県コード、市町村コードを用いることにより行政区域単位で所在地を表わすことができる。地理的な位置は経緯度法によるメッシュコードで表わすことができるが、これは別ファイル（Site Description File）でもつので測定局コードとは切り離す。②については測定局の管理者（国、都道府県、市、民間企業）の区別が表現できるものとする。③についてはその測定局が一般環境大気測定局か沿道大気測定局（自動車排出ガス測定局）かの区別が表現できるものとする。④については、地上からの高さに応じていくつかの測定がされている場合、その区別を表現できるものとする。

以上の考え方から、表 1 1 に示すような 1 0 桁コード（ $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}$ ）を考える<sup>10)</sup>。このように各測定局にコード与えることにより、各測定局を唯一のものとして特定することが可能となる。しかし、 $X_3, X_4$ は説明的なコードであるので測定局を特定するだけであれば不可

欠なものではない。

表 1 1 大気汚染測定局コードの設定法

大気汚染測定局のコード X <sub>1</sub> , X <sub>2</sub> , X <sub>3</sub> , X <sub>4</sub> , X <sub>5</sub> , X <sub>6</sub> , X <sub>7</sub> , X <sub>8</sub> , X <sub>9</sub> , X <sub>10</sub>	
X <sub>1</sub> , X <sub>2</sub>	都道府県コード（総務庁標準地域コードによる）
X <sub>3</sub> , X <sub>4</sub>	測定局管理者区分コード（都道府県：0 1、政令市：0 2、政令市以外の市町村：0 3、国：0 4、企業：0 5、不明：9 9
X <sub>5</sub> , X <sub>6</sub> , X <sub>7</sub>	市町村コード（総務庁標準地域コードによる）
X <sub>8</sub> , X <sub>9</sub>	測定局番号、市町村の中に存在する測定局に 1 番から付番する。ただし、0 1～5 0 は一般環境大気局に割当て、5 1～8 0 は自動車排ガス局に割当て、8 1～9 9 はその他の局に割当てる
X <sub>10</sub>	同一測定局において、垂直方向に複数か所サンプリングしている場合、サンプリングか所ごとに 1～9 の範囲で付番する。一般には 0 を与える。

b) 測定項目コード<sup>10)</sup>

標準コードとしてアルファベットで表現するコードと数字で表現するコードの 2 種類を考えた。前者は見ただけで内容が判断できる利点があり、後者はコンピュータによるデータ処理を容易かつ迅速にする利点がある。コードの桁数は 4 ですべて左づめで余白はスペースとする。標準コード案を表 1 2 に示す。

c) 単位コード<sup>10)</sup>

標準コードとしてアルファベットで表現するコードを採用する。桁数は 6 ですべて左づめで余白はスペースとする。たとえば ppb は P P B △ △ △、pphm P P H M △ △（△はスペースを意味する）などと表わす。測定項目ごとに特定の単位コードが指定されるので他の単位で表わすことはできない。



表 1 2 標準測定項目コード、単位コード案

区分 コード	測定項目	測定項目 アルファベット	測定項目 数字コード	測定値単位コード	測定値単位
大 気 汚 染 物 質	1. 二酸化硫黄	S O 2	0 0 0 1	P P B	ppb
	2. 窒素酸化物	N O X	0 0 0 2	P P B	ppb
	3. 一酸化窒素	N O	0 0 0 3	P P B	ppb
	4. 二酸化窒素	N O 2	0 0 0 4	P P B	ppb
	5. 一酸化炭素	C O	0 0 0 5	0 . 1 P P M	0.1 ppm
	6. オゾン	O 3	0 0 0 6	P P B	ppb
	7. オキシダント	O X	0 0 0 7	P P B	ppb
	8. 炭化水素(プロパン換算)	T H C P	0 0 2 1	P P H M	pphm
	9. 炭化水素(メタン換算)	T H C M	0 0 2 2	P P H M	pphm
	10. 炭化水素一飽和	S H C	0 0 2 3	P P H M	pphm
	11. 炭化水素一不飽和	U H C	0 0 2 4	P P H M	pphm
	12. 炭化水素一ノンメタン	N M H C	0 0 2 5	P P H M	pphm
	13. 炭化水素一メタン	C H 4	0 0 2 6	P P H M	pphm
	14. 非メタン炭化水素計による炭化水素	T H C	0 0 2 7	P P H M	pphm
	15. 浮遊粉塵(ディジタル)	S P	0 0 3 1	M G / H M 3	1/100 mg/m <sup>3</sup>
	16. 浮遊粒子状物質	S P M	0 0 3 2	M G / H M 3	1/100 mg/m <sup>3</sup>
気 象	17. 風向	W D	0 0 5 1	1 6 D I R C	16方位コード
	18. 風速	W S	0 0 5 2	0 . 1 M / S	0.1 m/S
	19. 湿度	T E M P	0 0 5 3	0 . 1 ▼ C	0.1 °C
	20. 湿度	H U M	0 0 5 4	P E R C N T	パーセント
	21. 日射量	S U N	0 0 5 5	C A L / C 2	cal/cm <sup>2</sup> /h
	22. 雨量	R A I N	0 0 5 6	M M	mm
	23. 紫外線	U V	0 0 5 7	C A L / C 2	cal/cm <sup>2</sup> /h
	24. 気圧	P R S	0 0 5 8	M B	ミリバール
	25. 交通量	C A R	0 0 7 1		カウント/h

## d) 用途地域コード

標準コードとして 0 1 から 0 9 までの数字で表現するコードを採用する。  
都市計画法<sup>14)</sup>による地域の用途区分(第 1 種住居専用地域、第 2 種住居専用地域、近隣商業地域、商業地域、準工業地域、工業専用地域)ごとにそれぞれ 0 1 から 0 8 の数字を、その他地域には 0 9 の数字を与える。

(4) 測定データの記録法<sup>10)</sup>

測定データの記録方法についての標準化を検討する。測定データの単位は前述のように測定項目ごとに特定の単位とし、データは整数とする。風向データについては、NNEから順に右回りに N まで 1 6 方位に 0 1 から 1 6 までの数字を与え、これにより表現する。カーム(無風)は 1 7 を与える。データは案 1 と案 2 はバイナリで表現し、案 3 は数字で表現する。データの記録順序は、たとえば案 1 の場合であると、4 月 1 日 1 時のデータから 1 年間のデータがタイ

ムシリーズに並んだ型式をとる。記憶エリアは、1か月分のエリアとして、どの月の分も31日分がとってあるので、たとえば4月31日（実際はありえない日）に相当する部分にもエリアがある。このエリアには特別の数字をあてはめておく。すなわち、上述のような測定値がありえないエリアには9998（これを無測定マークと称する。）をあてはめる。無測定マークがあてはめられるのはこのような場合のほか、年度の途中から測定が開始され、それ以前のデータがないような場合である。一方、測定値が欠測の場合については、これとは異なる特別の数字をあてはめることとし、この数字には9999を用いる。

#### (5) 標準フォーマット

以上の検討により標準磁気テープフォーマットとして次の3案を提案する<sup>10)</sup>。  
すなわち、

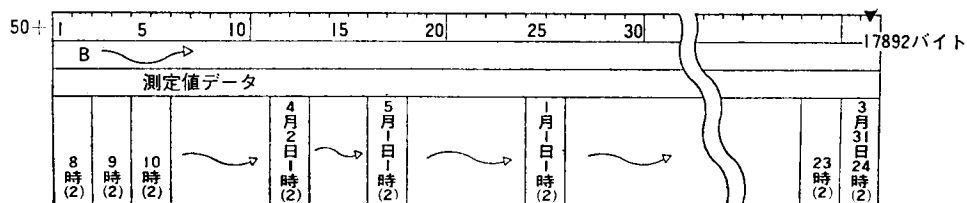
##### ① フォーマット案1

本フォーマットは、1レコード中に1測定局、1測定項目、1年間（4月から翌年3月まで）のデータ（ $24 \times 365 = 8760$ 個）を入力するものである。本フォーマットの特徴は、大型のコンピュータではデータ処理が能率的に行えること。大量のデータをコンパクトに蓄積できることである。また、論理構造も簡単でデータの読み込みも容易である。しかし、記憶容量が大きなコンピュータを必要とする点が短所である。実際にこのフォーマットでデータファイルを作成し<sup>15) 16)</sup>、運用することによりその実用性を確認したが、これについては第4章で述べる。レコードサイズは17892バイトである。測定局コードについては標準コードを入力できるエリアを設けると共に、各自治体で用いている自治体の独自コードも必要に応じて入力できるようエリアをとった。測定項目コードについても同じ理由から標準コードと自治体の独自コードのエリアをとった。本案のフォーマットを図3に示す。

##### ② フォーマット案2

本フォーマットは、自治体で用いられている（記憶容量がさほど大きくない）コンピュータでも適用できるよう、1レコード中に1測定局、1測定項目、1か月のデータ（ $24 \times 31 = 744$ 個）を入力するもので、レコード

バイト→ コード→ タイプ	1	5	10	15	20	25	30	35	40	45	50				
	E	E	E	E	E	E			E	B	B				
	西 暦  (4)	和 暦  (2)	測定項目 コード (標準 コード) (4)	同 左 (自治体 コード) (4)	単 位 名  (6)	用途地域 コード (2)	測定局番号 (標準)			測定局 番 号 (自治体 独自) (4)	測定データ (1時間値)				
							都道府県 コード (2)	管 理 者 区 分 コード (2)	市町村 コード (3)		測定局 コード (3)	4月1日 1時(2)	2時(2)	3時(2)	4時(2)



記録コード { E:EBCDIK      レコード長: 17892バイト/レコード(固定長)  
                  B:バイナリ      1レコード/1ブロック  
( ): バイト数

図3 標準磁気テープフォーマット案1

サイズは1550バイトであり、案1のそれよりかなり小さいものである。測定局コードおよび測定項目コードについては、案1と同様、標準コードと各自治体独自のコードを入力できるエリアがとってある。本案のフォーマットを図4に示す。

### ③ フォーマット案3

本フォーマットは、自治体のモニタリング用コンピュータでも適用できることを考慮して、レコードサイズを案2よりさらに小さくし、入力情報も簡略化している。すなわち、1レコード中に1測定局、1測定項目、1日のデータ(24個)を入力するもので、レコードサイズは132文字である。測定局コードについては標準コードと自治体独自のコードが入力できるエリアがとってあるが、測定項目コードは標準コード(数字)を用いることを前提にしている。また単位コードは標準単位表示がされていることを前提にしてこのためのエリアを設けない。これは繰り返しによる記録エリアの消費を少なくするためである。本案のフォーマットを図5に示す。



## 第5節 結 語

本章では大気環境データの組織化を図るための方策を検討した。まず、データの組織化の意義について論じ、ついでデータの観測、移送、チェック、編集・入力、蓄積・管理の各段階ごとにデータの組織化の課題を検討した。また、わが国の大気環境データ処理システムの現状分析をすることにより、各自治体における磁気テープのデータ記録方法が移種多様であることがデータの組織化の大きな支障であることを指摘し、これについての具体的な対応策——磁気テープのデータ記録方法の標準化——を検討した。対象となる事項の中で標準化することが可能で必要性も高い事項はハードウェアに依存しない事項、すなわちデータの配列(レコードフォーマット、レコードサイズなど)及び各種コード(測定局コード、測定項目コード、単位コード、欠測コードなど)の標準化であることを指摘し、データの配列の標準法として3案を、また各種コードについてもそれぞれの標準案を提案した。3案のうち1案は、レコードサイズが大きく能率的なデータ記録及び処理が可能なものでデータセンターにおけるデータファイル用として適するものであり、その他の2案は、レコードサイズが比較的小さく小型のコンピュータでも適用できるので自治体におけるデータファイル用として適しているものである。なお、提案した案(案1)は第4章で論じるデータベースのデータファイルの作成に実際に用いた。

## 第2章 水質及び廃棄物データの組織化

### 第1節 概 説

大気環境データの組織化についての前章の論議と同じ観点から、本章では水質環境データと廃棄物関係データについてデータの組織化を論じる。まず水質環境データの有する課題をデータの測定、チェック、集計・編集、蓄積の各段階ごとに検討する。そしてデータの蓄積段階の課題のうちデータの磁気テープへの記録方式については、このデータに適した方式を検討し提案する。廃棄物関係データについては、データの管理という点で、大気環境データや水質環境データほど進んでいないが、今後これに対する社会的必要性は大きくなる要素が大である。この分野の情報管理の現状をレビューし、今後の方向の考察を行う。

### 第2節 水質環境データの組織化

#### 2.1 データの測定及び処理の現状

我国における河川、湖沼、海域の水質監視測定システムは、環境庁、地方公共団体等関係行政機関により近年かなり整備されてきた。研究機関等においても、特定の水域ではその目的とするデータ把握のため独自の監視測定計画が立てられ実施されている<sup>1)</sup>また、国際的には、世界的な水質監視網を確立する目的で世界水質監視計画（GEMS/Water）がWHO（世界保健機構）とUNEP（国連環境計画）が中心となって進められている<sup>2)</sup>。

水質汚濁の監視・測定システムは、その目的、対象水域、測定機関、測定方法（サンプリング頻度、測定機器の種類）などの点で多様である。ここでは公共機関が法律等にもとづいて継続的に実施する水質監視測定として、次の3つのものをとりあげ、その現状についてレビューする。

##### ① 水質汚濁防止法に基づく水質監視測定システム

水質汚濁防止法<sup>3)</sup>により自治体等の水質監視の義務が規定されており、都道府県ごとに毎年測定計画を作成し、これによって水質の測定が行われている。

58年度には全国で人の健康に関する環境基準項目（カドミウム等9物質）の



測定は、5239測定点（河川3666、湖沼231、海域1342）で179,441検体が、生活環境保全に関する環境基準項目（BOD、COD等）の測定は、6811測定点（河川4232、湖沼378、海域2201）で414,477検体が調査された<sup>4)</sup>。これらの測定の実施は、1測定点において、月1回ないし月4回程度の頻度であることが通常であり、年間12～48回の測定が行われる。1回の測定は1サンプリングである。このような測定は通年調査と呼ばれている。これに対して、年間、1回～数回特定の日に、24時間の水質変動を経時的に調べようとする測定も行われることがある。これは通日調査と呼ばれ、通常、サンプリングは2時間おき、合計13回のサンプリングがなされる。通年調査のサンプリング頻度、通日調査の実施は各年度ごとに策定される測定計画<sup>5)</sup>にもとづいて定められるので年度によって、測定点によって異なることがある。測定項目は、環境基準の対象として取上げられている項目であるので、一般には表1に示すようなものであるが、測定点により、測定を省略されている項目もある。一方、これらの項目以外に特定の項目を追加して測定しているところもある。これらの測定結果は、自治体ごとにまとめられ、国（環境庁）に報告され、この結果をもとに毎年、環境庁監修「水質年鑑」として公表されている。記載されているデータは、全データ全てではなく、平均値及び $m/n$ （ $n$ は総検体数、 $m$ は水質環境基準に合致しない検体数）などである。

表1 水質汚濁に係る環境基準の項目

健康項目	生活環境項目		
	河川	湖沼	海
カドミウム	PH	PH	PH
シアン	BOD	COD	COD
有機燐	SS	SS	DO
鉛	DO	DO	大腸菌群数
クロム（6価）	大腸菌群数	大腸菌群数	n-ヘキサン抽出物質
ヒ素		T-P	
総水銀		T-N	
アルキル水銀			
P C B			

一方、昭和４５年ごろから重要な地点に水質の自動測定機が設けられ始めた。現在、全国で約２６０か所（建設省によるものを含む）に自動測定機が設けられ水質監視がなされている。東京都の例では、都内河川２９か所に「水質汚濁常時測定室」と呼ばれる測定局を設け、水温、ＰＨ、濁度、溶存酸素、電導度の５項目が全測定点で測定され、測定点によっては、その他塩素イオン、CODシアン等が測定されている。これらのデータは１０～２０秒おきにデータレコーダの上にプロットされ、１０分おきに中央局に有線テレメータで電送されている。中央局では、これを受信して、１時間平均値、日平均値、その他の統計値が集計されている<sup>6)</sup>。

## ② 建設省による監視測定システム

建設省では、河川管理の立場から主に１級河川について、月１～４回の頻度で水質測定を行っている。１回の測定では、原則として、６時間おきに４回のサンプリングをし、測定をする方法がとられている。現在、測定点の数は、全国で１０９水系１,０３１か所である<sup>7)</sup>。建設省でも、昭和４６年から自動測定機を主要地点に設けてデータ収集を行っている。その設置数は年々増加し、現在、全国６４水系の１２３個所に及んでいる。このうち４８水系９７か所の測定機にテレメータが接続され、工事事務所等で測定データの常時把握ができるようになっている。測定項目は、水温、ＰＨ、濁度、溶存酸素、導電率が全測定点で測定され、測定点によっては、その他シアン、アンモニア、COD、TOC等が測定されている<sup>8)</sup>。

## ③ 水道局による監視測定システム

各都市の水道局では、浄水操作の必要性から原水（河川水、湖沼水）の水質測定を行っている。原水の水質測定の頻度により、「毎日測定」、「週１回測定」、「月１回測定」などと呼ばれる原水の水質測定が行われている。どれを採用するかは、各浄水場の事情により異なっているが、全国の浄水場のうち８８％の浄水場では「毎日測定」を行っている<sup>9)</sup>。自動測定機による測定も行われており、全国２０５の浄水場で原水の水質自動測定機が設けられている<sup>9)</sup>。測定項目は、濁度、ＰＨ、水温、アルカリ度、導電率、シアン、溶存酸素、塩素

イオン等であり、自動測定機の数総計は653である<sup>9)</sup>。

## 2.2 データの組織化のための課題

コンピュータシステムの支援によるデータ管理を行い、データの総合的な利用を行おうとすると、その前段階として検討しなければならないデータ処理上のいくつかの問題点がある。これらは、前章で述べた大気環境データについての問題点と同様のものも多いが、水質環境データ特有のものであるものもある。以下、これらについて、データ測定段階、データチェック段階、データ集計段階、データ蓄積段階のそれぞれについて述べる。

- ① データ測定段階：各自治体等による公共用水域の水質汚濁の監視測定は法制度上の規定<sup>3)</sup>に基づいて行われているので大筋において全国的に統一された方法で行われていると言える。たとえば、各水質項目の分析方法は環境庁告示<sup>10)</sup>により規定されており、これに基づいて測定が実施されているのでこの点に関する限り問題はない。しかし、サンプリングの方法、(時、場所、頻度など)は法制度的に統一しきれない性質のものであり、測定計画<sup>5)</sup>に基づいて自治体ごとと測定点ごとにそれぞれの方法が採用されることとなる。この点はデータ処理の上では問題となる。例えばサンプリング頻度については、測定点により、月1～4回、つまり年間12～48回程度であり一様ではない。また各測定点におけるサンプリングの時刻は同一でないのが通常である。これは、ある水域における各測定点のサンプリングは通常同一日にするが、実際の作業上、当該水域の測定点について順次サンプリングしてゆくこととなるためである。サンプリング場所についても次のような点は留意が必要である。例えば、採水水深は原則として一定であるが、その目的により、いくつかのケースが設けられることがある。河川の場合、右岸、中央、左岸などでいくつかの測定がされることもある。以上のように水質環境データは形態の一樣でないデータであるが、これが本来的な特質であることに留意し、この組織化にあたっては、形態の統一を図る余り、データの有する情報量を減ずることのないよう配慮する必要がある。
- ② データチェック段階：水質環境データについても、多量のデータの処理を行

う場合には必ず異常値が出現し、この取扱いが問題である。マニュアル測定の場合には、自動測定機測定のように機械的に生じる異常値は原則として存在しない。異常値が混入する可能性は、分析誤差や分析上のミスによるものである。一方、降雨等の影響で水質が極端に汚濁したような場合の測定値についてはデータ処理上考慮が必要である。測定データのチェック方法としては第5章第3節で述べるような上下限值チェック法や測定点間比較チェック法などが考えられる。自動測定機測定のデータのチェックについては、大気環境データのチェックに用いられる方法の原理を採用しうるのであろう。

- ③ 集計・編集段階：ある水域の水質汚濁の状況の評価はその水域に係する測定点群のデータの総合的评价によってなされねばならないので、これが容易となるよう集計・編集がされる必要がある。そのため、測定点単位にまとめられたデータ群を水域単位のデータ群として検索、集計しうることが必要である。また、使用頻度の高い指標（ $m/n$ 、75%値、平均値、最大値など）を集計、編集したファイルを作成しておくと共に、これらの測定点ごと、水域ごとの経年変化が出力しうるようにしておくべきであろう。
- ④ 蓄積段階：データの総合的有効利用を目的としてデータをコンピュータシステムに蓄積する前提として、まずデータを磁気テープに入力することが必要である。データが全国的に共通的に利用しうるようにするためには全国の自治体で共通に使用可能なように標準的な磁気テープフォーマットを検討することが重要である。標準化すべき磁気テープフォーマットの具体的中味は、例えばレコードフォーマット、レコードサイズ、コード（測定点コード、測定項目コードなど）である。

### 2.3 データの蓄積方式の標準化

コンピュータシステムによるデータ管理を可能にするためには、前述のような問題点について具体的な検討を加えなければならない。このうち、以下では、データ蓄積段階の問題について検討する。具体的には、磁気テープへのデータの蓄積方式の標準化を考える。

水質環境データの場合は、大気環境データのように測定、送信、処理が自動化、機械化されていないものが大半であるので、各測定機関においてデータが磁気テープに記録され保管されることは多くない。しかし、今後、自動測定機が増えると考えられるし、マニュアル測定のデータでも、その量の増加のゆえに磁気テープによるデータ記録保管が普及するものと考えられる。したがって、各測定機関における磁気テープへのデータの記録保管の方式を標準化することが大気環境データと同様必要である。

水質環境データは、データが規則的なタイムシリーズとなって並んでいないのが通常であるという点（マニュアル測定データの場合）や、測定点の位置関係が自治体の領域内の位置関係より水系内におけるそれの方が重要である点など、大気環境データとは異なる点が少なくない。これらの点を考慮して、レコードフォーマットや測定点コードを考える必要がある。

前章 3.2 において述べたように、磁気テープの仕様の重要な部分は、①コンピュータの機種により必然的に決まるもの（記録密度、トラック数、記録コードなど）、②データの配列（レコードフォーマット、レコードサイズなど）、③各種コード（測定点コード、測定項目コード、単位コード、欠測コードなど）に分けられる。前章第 4 節で述べたと同じ理由により、標準化は②、③についてのみ検討する。各自治体で作成される磁気テープが、①の部分について標準化されないものとなっても、この部分は、コンピュータメーカ等によって既に作成されているコンバージョン・プログラムにより、比較的容易に変換が可能である。

磁気テープへのデータの蓄積方式については、マニュアル測定データを対象に考える。自動測定機による測定データは、その形態が大気環境データと類似であるので、前章で述べた蓄積方式が参考になりうる。データファイル中のデータが有すべき情報は、データそのものの他に、4W1H(WHAT, WHEN, WHERE, WHO, HOW)に関する属性情報である。これを具体的に水質環境データにあてはめると図 1 の上半分に示すようなデータとなる。データ処理操作上の利便性から、これらのデータを、同図に示すように、原データを入力する Raw Data File と集計値等を入力する Statistical Summaries File と測定点の位置の情報を入力す

る Site Description File の 3 種類とする(場合によっては, Statistical Summaries File と Site Description File は同一ファイルとしてもよい)。以下、Raw Data File を中心に論じる。

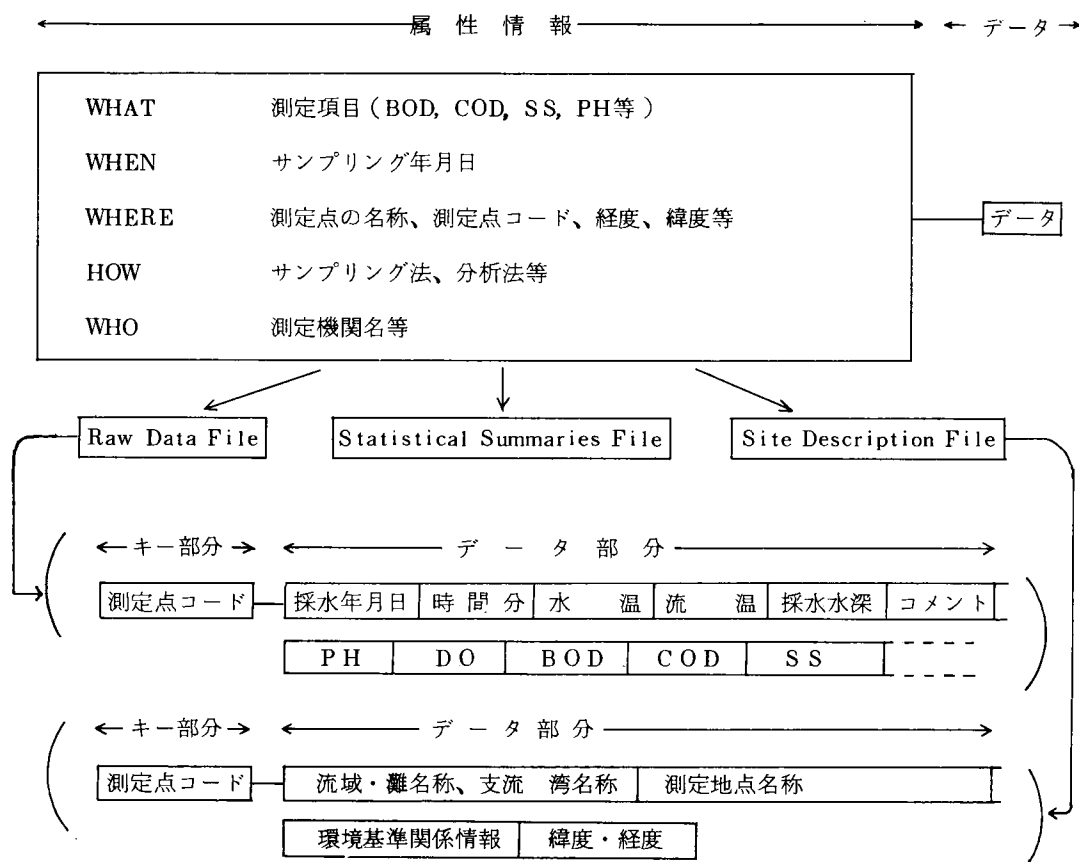


図1 水質環境データファイルの構成とフォーマットの概念図

#### (1) レコードフォーマットの検討

1レコードの中にデータそのものと4W1Hの属性情報を効率的に(同じ情報の繰り返しをできるだけ避けて)蓄積することを考える。水質の測定は、同じサンプリング試料で多くの項目を測定することが多いので、WHEN, WHEREの情報は同一で、WHAT(測定項目)ごとにデータがあるというデータ型式を考える。測定項目の種類と数は必ずしも一定でないが、考えられる項目を網羅し、項目の配列順序、位置を固定すると、格納される場所によって項目が何かを判別でき

るので、各データにWHATの情報を付加する必要がない。しかしこの方式では一部の項目しかデータがない場合でもスペースは同じだけ消費することになる。図1の下半分にこのような考え方によるレコードフォーマットの概念図を示している。

## (2) レコードサイズの検討

前述のようなレコードフォーマットを考えると、レコードのデータ部に入力されるデータは約80程度であり、レコードサイズは比較的小さい。大気環境データの場合はレコードサイズが非常に大きくなったので、中小型コンピュータでも利用できるよう、レコードサイズの小さい案も合わせて提案したが、この場合はその必要はない。

## (3) コード等の検討

属性情報をできるだけ標準的なコードで表現すると、データのコンピュータによる処理が容易になる。コードにより表現することが望ましい属性情報には、測定点、測定項目、測定状況などの情報がある。これらを表現すべく設定されるコードは前章第4節で論じた必要条件を満たすことが必要である。

測定点コード：測定点に関する情報として、具体的には次のような情報が必要である。

- ① 測定点の水域（河川か湖沼か海域か）
- ② 測定点の存在する都道府県
- ③ 測定点の存在する水系又は灘及びその支流又は湾
- ④ 測定点の番号

これらのうち、②と③はどちらか一方が必要である。①、②、④によってコードを構成したものと①、③、④によって構成したものを考え、前者をA案、後者をB案とした。A案は表2に示すような7桁のコード( $X_1, X_2, X_3, X_4, X_5, X_6, X_7$ )とし、B案は表3に示すような15桁のコード( $X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{13}, X_{14}, X_{15}$ )とした。A案はコードが都道府県単位に定められ、その変更があっても他の都道府県のそれに影響を及ぼさないという長所を有する。このコードでは都道府県単位のデータ処理は容易であ

るが、水域単位の水質処理の実施には不便である。一方、B案ではコードが水域単位に定められるので、同一水域に複数の都道府県が存在する場合、それぞれの測定点のコードは整合していなければならない、その変更は他に影響を及ぼす。しかし、水質環境データの大きな特徴として、データの解析は自治体の領域単位よりも水域単位に行う必要性の方が高いので、B案の方がデータ処理上の便利さを有する。しかしA案の取り扱い上の簡便さも捨てがたい。

表2 水質測定点コード（A案）

X <sub>1</sub> X <sub>2</sub> X <sub>3</sub> X <sub>4</sub> X <sub>5</sub> X <sub>6</sub> X <sub>7</sub>						
X <sub>1</sub> X <sub>2</sub>		都道府県コード（総務庁の標準地域コードによる） 0 1～4 7				
X <sub>3</sub>		水域分類コード（R, L, Sのいずれかとする） R：河川、L：湖沼、S：海域				
X <sub>4</sub> X <sub>5</sub> X <sub>6</sub> X <sub>7</sub>		地点番号0 0 0 1～9 9 9 9の番号を付ける。				

表3 水質測定点コード（B案）

X <sub>1</sub> X <sub>2</sub> X <sub>3</sub> X <sub>4</sub> X <sub>5</sub> X <sub>6</sub> X <sub>7</sub> X <sub>8</sub> X <sub>9</sub> X <sub>10</sub> X <sub>11</sub> X <sub>12</sub> X <sub>13</sub> X <sub>14</sub> X <sub>15</sub>														
X <sub>1</sub>		水域分類コード（1, 2, 3のいずれかとする） 1：河川、2：湖沼、3：海域												
X <sub>2</sub> X <sub>3</sub> X <sub>4</sub> X <sub>5</sub>		海区・離コード 河川及び湖沼では流出先の海区、離、海では当該海区、離につけられるコード X <sub>2</sub> X <sub>3</sub> ：大海区コード（たとえば太平洋南海区、東シナ海区など） X <sub>4</sub> X <sub>5</sub> ：離コード（たとえば大阪湾、東京湾など）												
X <sub>6</sub> X <sub>7</sub>		本川コード 海域に直接注ぐ支川（たとえば利根川、石狩川など）につけられるコード												
X <sub>8</sub> X <sub>9</sub>		一次支川コード 本川から分岐する支川（たとえば利根川の支川の渡良瀬川）につけられるコード												
X <sub>10</sub> X <sub>11</sub>		二次支川コード 一次支川から分岐する支川につけられるコード												
X <sub>12</sub> X <sub>13</sub>		三次支川コード 二次支川から分岐する支川につけられるコード												
X <sub>14</sub> X <sub>15</sub>		地点番号 三次支川内の地点番号												

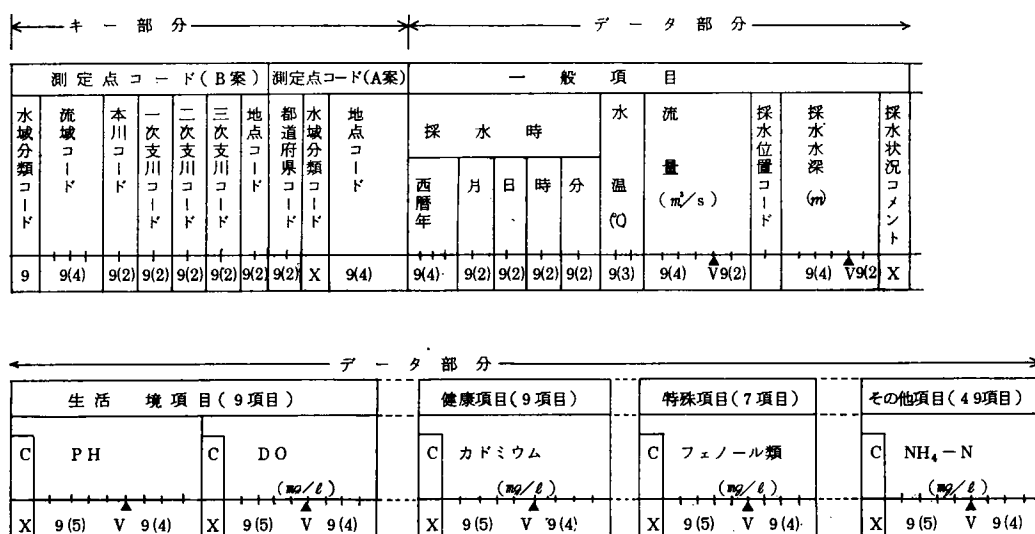


測定項目コード：水質環境データファイルのレコードフォーマットを前述のようなものにすると、各測定項目はそれぞれレコード中の所定の位置に格納されるので、このことにより、それぞれの測定項目が特定される。したがって、測定項目コードは用いない。

測定状況コード等：サンプリング時の気象や水域の状況のコメントを類型化してこれをコード化すること、設定されている環境基準をコード化すること、サンプリング位置（河川の場合は、右岸、左岸、中央、流心及びその組合せ、湖沼、海域の場合は、上層、下層、中層及びその組合せ）をコード化することなどを検討した。各コードの具体案の説明は割愛する。

#### (4) 標準フォーマット

以上の検討により、標準的磁気テープフォーマットとして考えたものを図2に示した。測定点コードはA案、B案の両方が併用される形をとっているが、これは、自治体内だけの利用を前提にデータファイルが作成される場合は、A



- (注) 9(4)は数字の4桁、9は数字の1桁を表わす。  
 X は文字の1桁を表わす。  
 V は小数点を表わす。▲は小数点の位置を表わす。

図2 水質環境データの標準的磁気テープフォーマット案

案によるコードがつけられると予想されることから、そのファイルが当該自治体以外の機関で利用されるとき、B案によるコードをつけ加えることにより、標準的な磁気テープ・フォーマットになるよう配慮したものである。フォーマットの中のデータ部分に格納する水質項目とその配列順序は表4に示すとおりである。

表4 磁気テープに入力される水質項目

生活環境項目	健康項目	特殊項目	そ の 他 項 目				
① PH	① カドミウム	① フェノール類	① $\text{NH}_4\text{-N}$	⑪ 導電率	②① Caイオン	②① TOD	④① 一般細菌
② DO	② シアン	② 銅	② $\text{NO}_2\text{-N}$	⑫ 透視度	②② Mgイオン	②② $\text{KMnO}_4$ 消費量	④② 腸球菌
③ BOD	③ 有機リン	③ 亜鉛	③ $\text{NO}_3\text{-N}$	⑬ 濁度	②③ 硫酸イオン	②③ 炭水化物量	④③ 植物プランクトン
④ COD	④ 鉛	④ 鉄	④ 無機性N	⑭ 蒸発残留物	②④ 酸度	②④ ABS	④④ クロロフィルa
⑤ SS	⑤ 6価クロム	⑤ マンガン	⑤ ケルダールN	⑮ 溶解性物質	②⑤ アルカリ度	②⑤ BHC	④⑤ 溶解性T-N
⑥ 大腸菌群数	⑥ ヒ素	⑥ クロム	⑥ 全有機性N	⑯ 塩分	②⑥ 全アルカリ度	②⑥ フタル酸エステル	④⑥ 浮遊性T-N
⑦ 油分	⑦ 総水銀	⑦ フッ素	⑦ アルブミニイドN	⑰ 塩素イオン	②⑦ ケイ酸	②⑦ ホルムアルデヒド	④⑦ COD(アルカリ法)
⑧ T-N	⑧ アルキル水銀		⑧ $\text{PO}_4$ イオン	⑱ 総硬度	②⑧ 硫化物	②⑧ アルミニウム	④⑧ COD(水道法)
⑨ T-P	⑨ PCB		⑨ 無機性P	⑲ Naイオン	②⑨ イオウ	②⑨ アンチモン	④⑨ 動物プランクトン
			⑩ 溶解性T-P	⑳ Kイオン	②⑩ TOC	②⑩ Ni	

### 第3節 廃棄物データの組織化

#### 3.1 データ処理の現状

廃棄物についてのデータは大気環境データや水質環境データと性格を大いに異にする。まず第一点は、データは、長い時間（例えば1年間）の広い空間（例えば市町村単位）における積算値又は平均値として表わされること、第2点は、廃棄物についてのデータの調査は法律によって義務づけられていないので、大気環境データや水質環境データと比べて、データ把握が定型化したものになりにくいこと。第3点は、同一項目のデータ量は多くないが、データの種類が多いこと。第4点は、他の自治体のデータを利用して総合的に解析しなければならない必要性が比較的小さいことなどがあげられる。

このような廃棄物データの特徴のため、現在まで、各自治体においてコンピュータを用いたデータ処理やデータの組織化は遅れていた。コンピュータを用いたデータ処理という点では、むしろ、国の方で全国的なデータの統計解析の際にその作業をコンピュータ化するという点の努力がなされ始めた。しかし、自治体における廃棄物データのコンピュータ化の試みも目芽えつつある。これについては次項で述べる。

各自治体において把握されている廃棄物データの現状については次のとおりである。廃棄物のうち一般廃棄物（法律<sup>11)</sup>の定義による）は、その処理が市町村の責務となっているので、市町村は、廃棄物の処理事業体として、事業実施に伴う統計資料の把握を行っている。この内容は一般廃棄物（ごみとし尿）の収集・運搬量、処理施設への搬入量、埋立処分量などである。これらのデータ以外にも関係データは数多く把握されているが、それは市町村（事業体）によって異っている。国（厚生省）では、年1回、一定の様式によってこれらのデータを調査、収集し、全国的な統計値を集計している。この結果は印刷され、行政資料として使われている<sup>12)</sup>。厚生省が調査しているデータ項目は各市町村の把握しているデータの共通的普遍的なものであるので、この調査項目を見ることにより、各市町村の廃棄物データの把握の実態が知られる。表5は厚生省の調査表の内容の概要を示したものである。集計は市町村・事務組合ごとのほか、表に示すように実施主

表5 一般廃棄物データの調査項目

データ項目	データ内容	集 計 単 位
ご み の 量	収 集 量	自治体・事務組合ごと、実施主体（直営、委託業者、許可業者）ごと、ごみの種類（混合ごみ、可燃ごみ、不燃ごみ、資源ごみ等）ごと
	処 理 量	自治体・事務組合ごと、処理方式（焼却、コンポスト、堆肥・飼料、埋立、その他）ごと
し 尿 の 量	収 集 量	自治体・事務組合ごと、実施主体（直営、委託業者、許可業者）ごと、し尿の種類（生し尿、し尿浄化槽汚泥）ごと
	処 理 量	自治体・事務組合ごと、処理方式（し尿処理施設、下水道投入、海洋投入、農村還元、その他）ごと
廃棄物の処理施設	ごみ処理施設の 数の能力  し尿処理施設の 数と能力  最終処分施設の 数と面積	自治体・事務組合ごと、施設の種類ごと
そ の 他	廃棄物処理業者 — 業者数、従業員数、機材  処理対象人口 — ごみ処理対象人口、し尿処理対象人口  処理事業経費 — 歳入内訳、歳出内訳  処理事業職員数 — 担当業務別	

体ごと、処理方式ごとにもまとめられる。一方、産業廃棄物（法律<sup>11)</sup>の定義による）については、その処理が事業者の責務となっており、その規制、指導の役割は都道府県（又は保健所設置市）にあるので、関係のデータの調査は、原則として、個々の事業場の有するデータを都道府県（又は保健所設置市）が集計して

いる。産業廃棄物のデータの把握は、法律による届出や報告徴収により行うことが可能なものと、別途にアンケート調査の方式で調査しなければならないものがある。処理施設や処理業者等についてのデータは前者に該当し、産業廃棄物の排出量や処理量等についてのデータは後者に該当する。前者に関するデータのうち特定のものについては、国（厚生省）が、年に1回、一定の様式により全国の自治体からデータを収集し、全国的な統計値を求めている。この結果は印刷され、行政資料として使われている<sup>13)</sup>。国による全国調査の対象となっているデータ項目は表6に示すようなものである。後者に関するデータについては、産業廃棄物の種類、業種ごとの排出量、中間処理量、資源化量、最終処分量などがこれに該当するが、大企業については個々の報告徴収や立入検査により調査し、中小企業などで実態の把握が十分でないものについては排出原単位を用いた推計方法により調査が行われている。しかし、排出原単位のとり方や業種のグルーピング等の方法は、必ずしも全国的に統一した方法で行われているとはいえない。

表6 産業廃棄物データの調査項目

データ項目	データ内容	集計単位
産業廃棄物行政組織	組織、職員数 公共関与機関の内容 行政処分の内容と件数	自治体ごと
産業廃棄物処理業者	許可件数	自治体ごと 許可対象事業ごと
産業廃棄物処理施設	中間処理施設の数 同 能力 最終処分施設の数 同 面積	自治体ごと 施設の種類ごと
産業廃棄物の処理実績	排出量 中間処理量 最終処分量 海洋投入処分量	自治体ごと 処理実施者の区分（事業者、処理業者、公共機関）ごと

以上のように、市町村や都道府県が調査、把握した一般廃棄物や産業廃棄物のデータの処理は大部分人手による作業で行われている。そして国においてもこれらのデータを収集して、全国的統計資料を作成する場合人手による集計作業によっていた。しかし、これは膨大な作業量であると共に集計ミスが生じるなどの問題があるので、厚生省（この時期筆者ら担当）では一般廃棄物のデータは昭和55年度から、産業廃棄物のデータは56年度からコンピュータを用いるデータ処理方式に改めた。この方式導入にあたっては、学識経験者や実務者からなる検討委員会を設け、システム設計を行った。例えば調査データのコーディング、パンチングの作業を合理化するため、一般廃棄物データでは回収した調査票のデータをコーディングすることなくそのままパンチングできるよう調査票の設計を工夫した。また、処理施設のコード等についても統一的な設定方法を示した。データは磁気テープに収録されるが、これをコンピュータシステムに入力し、処理し、一定様式に出力したものの一部を図3に示す。

### 3.2 情報管理の方向

前述したように、各自治体の廃棄物関係データは、コンピュータによる総合的管理が可能な状態に至っていない。しかし、今後、廃棄物対策を科学的に行うためにはこれは不可避の課題である。自治体における廃棄物の情報管理の必要性は、主として次のような点に存在すると考えられる<sup>14)</sup>。

第一は、廃棄物（とくに産業廃棄物）の規制に関する行政事務を迅速かつ科学的に処理すること。たとえば、廃棄物排出事業所等からの届出書類、報告書類、廃棄物処理業の許可に関する書類などの情報管理などである。発生する廃棄物の質と量についての情報管理も当然含まれる。第二は、中間処理施設や最終処分場へ搬入される廃棄物（一般廃棄物及び産業廃棄物）の質及び量を管理すること。たとえば、搬入される廃棄物の質のチェック、量の計量と集計、料金の算定と請求などに関する情報管理である。第三は、事業活動及び社会活動と廃棄物の発生と関係の予測、すなわち生産プロセスや事業活動に応じて発生する廃棄物の種類と量及び性状を推定したり、またその廃棄物の物流状況を推定することなどによ

[illegible]

### ③ 廃棄物データのコンピュータ処理結果の出力例

り、これらが環境に対してどのような影響を与えるかを予測するための支援システムとしての情報管理である。

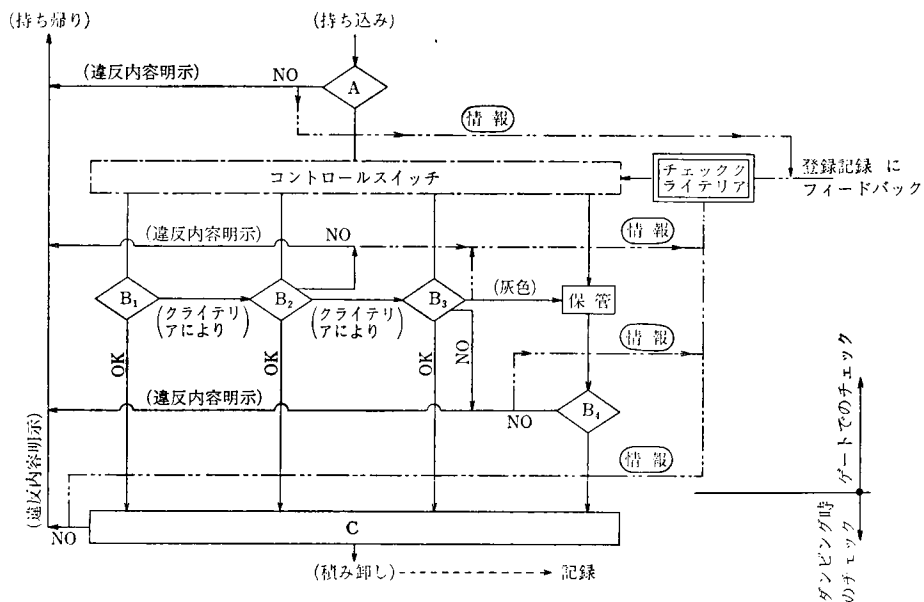
自治体における廃棄物の情報管理は、第一のものから手をつけられ、逐次第二、第三のものも検討してゆくといった発展過程を踏むものと考えられる。すでに一部の先進の自治体ではこれらのシステムを作成（計画中を含む）しているところもある。以下ではこれらについて述べる。

第一のものについてはすでに横浜市で採用され、実務に用いられている例がある。この例では、産業廃棄物の規制に関する情報が対象になっており、例えば、管内の許可業者に関する情報（名称、所在地、取扱い廃棄物の種類と量、処理方法など）がコンピュータ（オフィス・コンピュータと呼ばれる小型のコンピュータ）に入力、蓄積されている。端末機から指示をすることにより一定の様式で希望する結果が出力される。例えば「〇〇地域に所在」し、「取扱う廃棄物の種類は〇〇」で、「その処理方法は〇〇」というのに該当する産廃処理業者は？というような質問に対して即座に回答が出力される<sup>15)</sup>。発生する廃棄物の質と量の情報管理に関連して、いわゆる廃棄物交換を支援する情報システムがある。このシステムは、ある事業所で発生した不要な廃棄物でも他の事業所では有効に利用しうる場合があるので、提供希望廃棄物に関する情報と入手希望廃棄物に関する情報を管理することにより廃棄物の再利用を促進させることを目的とするものである。情報量によってはコンピュータ支援によるシステムを考えることは必ずしも必要ではない。大分県、山口県、愛媛県では、定期的に関係の情報を印刷物の形でユーザの目にふれるようにしておく方式で、すでに実際に運営されている<sup>16)</sup>。

第二のものについては、大阪湾広域臨海環境整備センター<sup>17)</sup>が検討している情報管理システムがあげられる。同センターは近畿2府4県の関係市町村で発生する廃棄物（一般廃棄物、産業廃棄物）の最終処分のための用地を確保するための広域的な事業である<sup>18)</sup>。この事業では、最終処分場に搬入される廃棄物の種類の判別、質のチェック、量の把握とその情報管理、それらの情報に基づく料金徴収のための情報管理などが必要であり、このため情報管理の機能として、①契約管理情報システム、②契約審査情報システム、③受入管理情報システム、④受入検



査情報システム、⑤料金徴収情報システム、⑥緊急時情報システムが考えられている<sup>19)</sup>。その中で、受入検査情報システムの概念は、筆者もメンバーである研究グループが提案したシステム（図4参照）<sup>20)</sup>の概念と同様である。



- A: ゲートにおいて書類照合、搬入者識別、秤量などを行う。  
 B: チェックライテリアにより搬入者、排出業種、廃棄物毎のチェックを行う。  
 B<sub>1</sub>: 目視による全体的検査、特に疑わしいところがなくチェックライテリアで B<sub>2</sub> 以降の検査が不要の場合は C へ、それ以外は B<sub>2</sub> のチェックを受ける。  
 B<sub>2</sub>: 水分状況、臭気、色など感覚的チェックを排出者提出のサンプルとの比較又は経験上から行う。契約の状況との相違がはなはだしいときは持ち帰り、そうでなく又チェックライテリアに B<sub>2</sub> 検査の指示されないものは C に、その後は B<sub>3</sub> チェックに回される。  
 B<sub>3</sub>: 新迅速分析法による pH, Hg, Cd, Pb, As, Cr, CN 等のチェック。分析は簡便法なので白又は黒と断定しうるときのほか、どちらとも断定しがたい灰色の結果のときもある。白であれば C に、黒であれば持ち帰り、灰色のときは 1 時保管して B<sub>4</sub> 検査を行う。1 時保管は一応は受入れるが B<sub>4</sub> 検査で黒と出れば搬入者を呼んで後日持ち帰えらせる。  
 B<sub>4</sub>: 環境庁告示 13 号による正規分析で 1 時保管後 1~2 日後でないとも結果が判明しない。  
 C: 埋立のダンピング時、基地ではホッパ投入時、積荷の中に混入しているものを目視チェック。

図4 廃棄物の受入検査情報システム<sup>20)</sup>

第三のものについては、計画策定または政策決定の支援システムとしてのデータベースが該当する。これは大気汚染関係データ、水質汚濁関係データ等も含めたいわゆる環境データベースの一部として考えるべきものであり、神奈川県、兵庫県、京都市等では地域環境管理計画の策定のための支援システムとしてこのよ

うなシステムの作成が検討されている<sup>21)</sup>。

なお、将来の方向としては、一自治体内で完結するシステムから発展させて、情報が他の自治体または国のそれと互換性があるようなシステムとして考えるべきである。なぜなら、産業廃棄物は発生した場所から処理、処分される場所まで複数の自治体にわたり広域的に運搬されることもあるので、適切な情報の相互交換を行い、それを的確な規制、指導に役立てる必要があるからである。このような観点から、将来自治体が情報管理システムを導入する際の指針をまとめる作業が国において進められている<sup>22)</sup>。また、情報管理の広域化の基本構想を確立する意味から近畿2府4県を対象に広域的な情報管理のケーススタディが行われている<sup>23)</sup>。

全体としてみた場合、廃棄物に関する情報管理の試みはまだ緒についたばかりと言うべきで、廃棄物の組織化をはじめとして今後に残された課題は多い。

#### 第4節 結 語

本章では、水質環境データの組織化と廃棄物関係データの組織化について検討した。以下に、検討結果を要約して述べる。

水質環境データについて

- ① 水質環境データの測定は、自治体、建設省、各都市水道局等で、それぞれの目的で、実施されている。この実態をレビューした。
- ② 自治体によってとられている水質環境データに関し、データ測定、データチェック、データ集計・編集、データ蓄積の各段階における課題を考察し、とくに、コンピュータによるデータ管理と利用を計画するには、データの蓄積方式の標準化が不可欠であることを述べた。
- ③ 磁気テープへのデータの記録方式標準化の対象となる事項のうち、レコードフォーマット、レコードサイズ、各種コードについて、標準化の具体的な案を提案した。
- ④ 標準的磁気テープ・フォーマットは、データが大気環境データのように規則的なタイムシリーズのデータではない点の特徴を考慮して、具体案を提案

した。提案した案は、第4章で論じるデータベースのデータ・ファイルの作成に実際に用いた。

廃棄物関係データについて

- ① 市町村で把握されている一般廃棄物関係データの現状及び都道府県等で把握されている産業廃棄物関係データの処理の現状をレビューした。
- ② 廃棄物関係データの情報管理は遅れており、今後の課題であることを指摘し、今後のこの分野における情報管理システムの機能を考察した。

### 第3章 国内外における環境情報システムの現状

#### 第1節 概 説

データの管理の方法は、データの記録、蓄積媒体の型式によって、①調表の形で管理、保存するもの、②コンピュータ・リーダブルな形（たとえば磁気テープ）で管理、保存するもの、③データをオンラインで利用できるようコンピュータシステム内に蓄積しておくものに分類することができる。データを総合的に高度に利用しようとする②、③の方法によるデータ管理を考えることが必要である。

米国では、大気環境データベース、水質環境データベースとして、それぞれ、SAROAD、STORETと呼ばれるものが設けられており、全国の測定データはこのシステムの中に入力、蓄積されている。ユーザはこのデータを端末器を通して入手することができるほか、備えつけのアプリケーション・プログラムを用いて各種の高度の解析を迅速に実施することができるようになっている。今後、わが国においても、このようなシステムが設けられ、環境データが便利に利用されることが可能になれば、関係の研究や施策の実施に大いに役立つであろう。そこで、環境データベースの作成の検討（次章）に先立ち、本章では、まずこの問題についてのわが国の状況を概観したのち先進諸外国におけるこれらの開発、運営の状況についてレビューする。

#### 第2節 わが国における環境情報システムの現状

情報システムが研究対象として組織的にとりあげられ、研究活動の総合化が促進され始めたのは文部省の特定研究「広域大量情報の高次処理」、「情報システムの形成過程と学術情報の組織化」が開始された昭和48～50年頃からである。「情報システムの形成過程と学術情報の組織化」では、入出力システム開発委員会、データベースマネジメントシステム開発委員会、ネットワーク開発委員会、学術情報システム開発委員会が設けられ、主としてシステムの開発に関する基礎的な研究が実施された。一方、実用的なデータベースの作成、運営も並行して進められて来た。しかし、全国の大学や研究機関で学術情報のデータベースがどの

ように開発されているかについての実態調査<sup>1)</sup>——前述の学術情報システム開発委員会（筆者参画）による調査——によると、学術情報についてのデータベースの作成は、物理学、化学、医学、薬学等の基礎科学分野のデータ（とくにいわゆる **fact data** と呼ばれるもの）において進んでいるが、他の分野では、まだその作成に手がつけられていないか、または外国の情報システムやデータファイルの購入によって対応している場合が多いとしている。

環境データを対象とした情報システムの開発、運営の例はさらに限られたものとなる。以下においてこの状況について概観する。

#### (1) 数値データの情報システム

環境に関する数値データでは環境汚染の測定データが最も代表的である。自動連続測定機による大気汚染の測定、テレメータシステムによるデータ送信、コンピュータによるデータ処理を組合わせた大気環境データシステムが多くの自治体に普及して来ている。このシステムは、汚染質や気象を測定する自動連続測定機を中心とした測定系、測定系からの出力情報を伝送路をとおして中央局に送るテレメータ系およびこれによって集められた情報に各種の演算処理をほどこすデータ処理系の3つに分けられる。<sup>2)3)</sup>データ処理系は中央処理装置と各種周辺装置とからなっており、テレメータ系で収集した各測定局ごとの情報から時間補正とスケール変換処理を行い、1時間値を算出し、各種表示盤へのデータ表示、時報印字、磁気テープへのデータの記録などを行う。このシステムは汚染の常時監視が目的のシステムであるので、3つの系のうち、測定系、テレメータ系に重点があり、データ処理の点は全体としては補足的な位置づけである。

一方、昭和50年頃から、汎用型大型コンピュータを用いた環境データを対象とする情報システムの採用が検討され始めた。これは環境アセスメントの実施や環境管理等の行政施策の実行のための情報提供の支援システムとして考えられている。このような目的の環境データベースは、実験的に作成されたもの<sup>4)</sup>を除き、現在までのところ自治体レベルで実用的に用いられているものはない。しかし、先進のいくつかの自治体でその開発・作成について検討がされている。そのねらいとするところは、「環境および発生源の監視、予測、制御に必要な

情報を収集、整備し、それを一元的に統合し、蓄積、管理することによって環境全体を把握し、適時に適切な情報を加工、提供することによって、総合的な地域環境管理行政の遂行を支援する。<sup>5)</sup>』というものであることが多い。

## (2) 文献情報システム

環境に関する文献についての情報システムとしてまず第一にあげられるのは、日本科学技術情報センター（JICST）で発行している科学技術文献速報（環境公害編）とその情報システムである。JICSTでは、内外の文献を抄録した日本語による抄録誌である「科学技術文献速報」を発行している。この中の一つである環境公害編は大気汚染、水質汚濁、廃棄物その他環境公害問題に関して発表された文献が収録されている。科学文献速報のうち文献のタイトル、著者、雑誌名、発行年、キーワードは磁気テープ（JICST文献検索用磁気テープ）に入力されており、ユーザは検索サービスを受けることにより情報を得ることもできるし、磁気テープそのものを購入し自分で検索、利用することも可能である。この磁気テープを用いて検索した例<sup>6)</sup>を図1に示す。これは、「大気汚染」と「環境保全」という2つのキーワードを有するものを検索した結果（106件のRESULT）の一部である。

このほか、外国の情報機関からファイルやシステムを購入したり、借りたりして文献検索サービスを行っている例もある。この代表例として、CAC 化学文献ファイル、MEDLARS 医学文献ファイルがある。環境関係では、Environmental Information Center Inc. のEIC 文献ファイルや米国NTIS（National Technical Information Service）のGRA（Government Reports Announcement）文献ファイルなどがある。

検索結果	国名	年次	言語	分野	キーワード	著者	タイトル
RESULT		106					*** NIES BUNKEN (KOGAI) ***
K75010030	USA	73	EN	P999A	P B	* BLEECKER H	CAPITAL AND OPERATING COSTS OF POLLUTION CONT
K75010377	DEU	74	DE	E590A	NEUE DELIWA Z	N0:8 SASSE W	PROBLEME DER ENERGIEVERSORGUNG IN DER STADTRE
K75020029	USA	N0	EN	D812A	PROC ANNU ENVIRO	VL:10TH WILLIAMS E	ENVIRONMENTAL LEGISLATION IN THE 1971 TENNESS
K75030526	DEU	74	DE	C084A	GLUCKAUF	VL:110 N0:2 HOTZEL E	VERWENDUNG VON STEINKOEHLE UND LUFTREINHALTUNG
K75030855	DEU	74	DE	B059A	DEUT BAUTZG	VL:108 N0:9 *	MULLVERBRENNUNG: - EIN UMWELTFREUNDLICHES VER
K75040037	SUN	74	RU	R002A	STR0IT ARKHIT LE	VL:36 N0:10 ZHURBA V T	OSTOR0ZHNO: PRIRODA UYAZVIMA.
K75040570	SCH	74	DE	B274A	CHIMIA	VL:28 N0:5 BOHLEN B	KONTROLL- UND UBERWACHUNGSAUFGABEN AUF DEM GE
K75040597	SCH	74	DE	B106C	GESUNDHEITSTECHN	VL:8 N0:11 SCHAETZLE P	DIE LUFTVERUNREINIGUNG IN DER SCHWEIZ.
K75040899	GBR	74	EN	D393A	PROC R SOC SER A	VL:339 N0:1 GOODMAN G T	ECOLOGY AND THE PROBLEMS OF REHABILITATING WA
K75040968	SCH	74	DE	B274A	CHIMIA	VL:28 N0:5 DEGISCHER G	UMWELTSCHUTZ DURCH VERBESSERTE INFRASTRUKTUR.
K75050011	JPN	74	JA	F228A	NIPPON KIKAI GAK	VL:77 N0:67 ヲカワ	カンキョウセツンキョシユツ ノ ゲンシヨクト シヨクタイ
K75050526	FRA	74	FR	D047A	REV METALL	VL:71 N0:11 GENY P	LA SIDERURGIE ET SON ENVIRONNEMENT.

図1 J I C S T ファイルによる文献検索例

### (3) 情報源情報システム

情報源に関する情報（欲しい情報を誰がもっているかを示す情報）のシステムは、国立公害研究所が昭和50年に作成したもの<sup>7)</sup>があるが、システム作成を主眼とした実験的なものであり、情報量の点では十分でない。一方、昭和50年から国連環境計画（UNEP）によって国際環境情報源照会システム（INFOTERRA と呼ばれている）の運営が開始された。<sup>8)</sup>このシステムは、世界各国の環境に関する情報を有する情報源を登録しておき——昭和59年現在、88か国の約9600機関が情報源として登録されている<sup>9)</sup>——ユーザに対しリファラルサービスにより、必要な情報源に関する情報を提供するというものである。国立公害研究所が国内フォーカルポイントに指定され、国内のこの活動を実行している。このシステムの機能の概略は、図2に示すように、まず

①求める情報がどこにあるかをフォーカルポイントに問合せ、②該当する情報源の紹介を受け、③当該情報源に情報提供を依頼し、④情報を得るというものである。

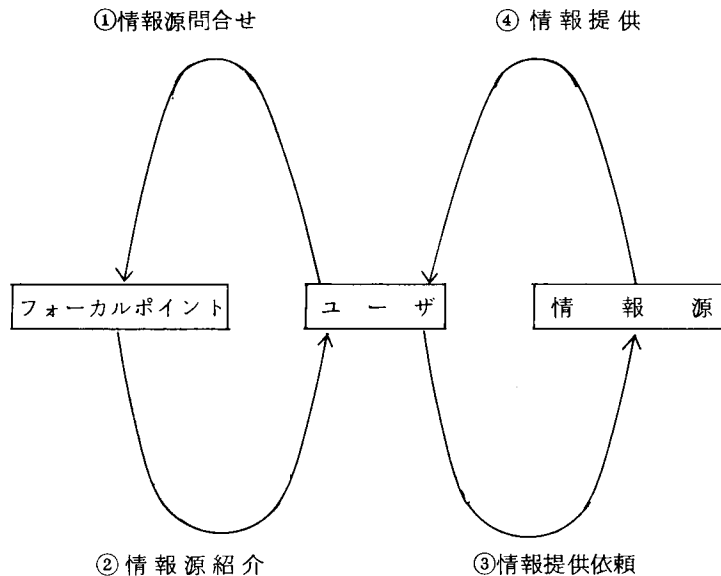


図2 INFOTERRAシステムの概要

### 第3節 諸外国における環境情報システムの現状<sup>10) 11) 12)</sup>

諸外国における環境情報システムの開発、運営状況を調べ、とくに先進的な例として、米国、カナダ、西独、スウェーデンの状況をレビューする。

#### 3.1 米 国<sup>13) 14)</sup>

米国においては、すでに1958年に大気環境データについてのデータベース、1964年には水質環境データについてのデータベースが作成されている。現在においては、その他の環境データについても多くのデータベースが政府機関（環境保護庁、農務省、商務省、内務省、エネルギー庁）において作成、運営されている。その種類は非常に豊富である。表1は関係のデータベースを分類して一覧表としたものである。また表2～表6は、それぞれ大気、水質、自然・資源・エ



エネルギー、化学物質、法律・予算に関するデータベースについて、表7は文献情報及び情報源情報に関するデータベースについて名称と内容を一覧表にしたものである。これらのうち代表的なものについて説明を加える。

表1 米国における主な環境情報システムのカテゴリ別一覧

	数 値 デ ー タ	文 献 情 報	情 報 源 情 報	作図, 作表, シミュレーション・モデル
大 気 汚 染	SAROAD, NEDS, CDS	APTIC		PRMS, RAPS, IPP
水 質 汚 濁	STORET, PCS, WATSTORE		NAWDEX	GLWQM
自然環境, 資源エネルギー等	CNI, SLPA, EROS	OASIS	ROSCOP	SEAS, SOS
騒 音		NOISE FILE		
廃 棄 物		SWIRS		
化 学 物 質	PARCS, ESPS, TAD			Predictive Model for Freshwater Eco-system
法律, 予算, その他	SRIS, GICS, FFF			

表2 大気関係データの情報システム

名 称	略 称	内 容	作成省庁
Aerometric and Emissions Reporting System	AEROS	SAROAD と NEDS のデータを統合したレポーティングシステム	(1)
Storage and Retrieval of Aerometric Data	SAROAD	全国の大気環境データのファイル	(1)
National Emissions Data System	NEDS	全国の大気汚染発生源のデータ (ばい煙発生量, 操業状態) のファイル	(1)
Plans Review Management System	PRMS	AEROS のデータを用いた作表, 作図プログラム	(1)
EPA Energy Data System	EDS	NEDS, SAROAD, CDS, FORM-67 を統合したファイル (作成中)	(1)
Form 67 Retrieval System		全国 825 発電所についての関連データのファイル	(1)
Compliance Data System	CDS	州ごとの大気汚染削減計画による発生源の届出事項のデータファイル	(1)
Emissions Control Data Base System	ECDBS	自動車排出ガスのテスト結果のデータファイル	(1)
Fuels Data Base System	—	無鉛ガソリンの検査による違反状況データのファイル	(1)
Air Quality Implementation Planning Program	IPP	大気汚染削減計画を作成するためのモデルプログラム	(1)
Regional Air Pollution Study	RAPS	シミュレーションモデルによる大気汚染の予測	(1)
Population Studies System	—	特定の大気汚染物質の人体影響を評価する統計データを作成	(1)
Air Quality Data	AIRQ	主要な5つの大気汚染物質 (EPA から提供) と気象データのファイル (作成中)	(5)

- (注) (1) U.S. Environmental Protection Agency  
(2) U.S. Department of Commerce  
(3) U.S. Department of Interior  
(4) U.S. Department of Agriculture  
(5) U.S. Energy Research and Development Administration  
(6) U.S. Department of the Army Corps of Engineering  
表3～表7も同様

表3 水質関係データの情報システム

名 称	略 称	内 容	作成省庁
Storage and Retrieval for Water Quality Data	STORET	全国の水質環境データのファイル	(1)
Permit Compliance System	PCS	水質汚濁削減計画にもとづく発生源の排出量、濃度等の届出事項のデータファイル	(1)
National Eutrophication Study	NES	全国800湖沼の水質と水流のデータ	(1)
Inventory of Public Water Supplies	INVWAS	全国における飲料用水の調査結果データ	(1)
Great Lake Water Quality Model	GLWQM	湖沼における排水負荷の削減効果の予測モデル	(1)
Predictive Model for Freshwater Ecosystem	—	殺虫剤、水銀、藻類の挙動と分解に関する予測モデル	(1)
Hydrologic Data Bank	HDB	水域ごとの雨量、流量、水温、蒸発量、地熱等についてのデータファイル	(4)
Water Data Storage and Retrieval System	WAT-STORE	陸水、地下水に関する水質、水量、測定点に関する情報についてのファイル	(3)

表4 自然環境、資源、エネルギー等に関する情報システム

名 称	略 称	内 容	作成省庁
State of the System Model	SOS	環境、資源の保全と人口増加の調和を図るための各種手法モデル	(1)
Strategic Environmental Assessment System	SEAS	経済、エネルギー、環境に関する主要なすべての因子による環境と相互作用のシミュレーション	(1)
Conservation Needs Inventory	CNI	土壌、土地利用、土地の保全措置に関するデータファイル	(4)
Environmental Data Index	ENDEX	下の5つのデータベースを傘下におさめる。各データベースに関する情報をファイル化	(2)
Environmental Data Base Directory	EDBD	全国の環境データベースの Directory	(2)
Earthquake Data File	EDF	地震、振動についてのデータベース	(2)
NODC Index for Instrument-Measured Sub-surface Current Observations	NIMSCO	海流に関するデータベース	(2)
Bottom Photography Camera Station File	PHOTO	海底写真撮映の時、場所等の情報のデータベース	(2)
Report of Observation/Samples Collected by Oceanographic Programs	ROSCOP	海洋関係の研究機関、研究者および研究内容についてのデータベース	(2)
Environmental Mutagens Data Base	EMDB	環境中の化学物質の突然変異テストに関するデータベース	(5)
Selected Legally Protected Animals	SLPA	要保護の野生動物に関するデータベース	(6)
Sensitive Wildlife Information System	SWIS	野生動物の生息、生態に関するデータベース	(6)
Environmental Satellite Data	ESD	航空写真および衛星写真情報のデータベース	(2)
Earth Resources Observation System	EROS	NASAのランドサット衛星写真情報のデータベース	(3)

表5 化学物質に関する情報システム

名 称	略 称	内 容	作成省庁
Pesticides Analysis Retrieval and Control System	PARCS	殺虫剤の製造者、成分、利用等についてのデータベース	(1)
Pesticides Registration System	—	殺虫剤の名称、化学組成、製造工場の名称と位置等についてのデータベース	(1)
Epidemiological Studies Program System	ESPS	殺虫剤の毒性、使用状況の調査結果についてのデータベース	(1)
Establishment Registration Support System	ERSS	殺虫剤の製造工場、製造量に関するデータベース	(1)
Pesticides Enforcement Monitoring System	PEMS	殺虫剤に関する規制、違反についてのデータベース	(1)
Technical Assistance Data System	TADS	863 種の有害物質の特性についてのデータベース	(1)

表6 法律、予算その他に関する情報システム

名 称	略 称	内 容	作成省庁
Standards & Regulations Information System	SRIS	EPA の排出基準、ガイドラインに関するデータベース	(1)
Grants Information and Control System	GICS	公害防止施設の国庫費助成の申請、採択、建設等に関するデータベース	(1)
Federal Facility Budgetary Data System	FFF	国家機関により建設される公害防止施設の予算についてのデータベース	(1)
Laboratory Data Management System	LDMS	EPA の研究所でとられたデータの蓄積、編集処理、解析システム	(1)

表7 文献情報、情報源情報に関する情報システム

名 称	略 称	内 容	作成省庁
Air Pollution Technical Information Center	APTIC	大気汚染の技術情報に関する文献システム	(1)
Solid Waste Information Retrieval System	SWIRS	固型廃棄物に関する文献システム	(1)
Noise File	—	騒音に関する文献システム	(1)
Oceanic and Atmospheric Scientific Information System	OASIS	海洋資源、環境科学研究の文献システム	(1)
Library Management System	LIMS	EPA の 28 の図最館における本、雑誌の管理および検索のシステム	(1)
National Water Data Exchange	NAWDEx	水関係データの保有機関、データ測定点に関する情報についての検索システム	(3)

## (1) 大気関係(表2)

代表的なものとして SAROAD, NEDS, AEROS と称されるものがある。

SAROAD(Storage and Retrieval of Aerometric Data)は全国の大気環境データのデータベース、NEDS(National Emission Data System)は全

国の大気汚染発生源データのデータベースである。この2つを結合させる目的でAEROS (Aerometric and Emissions Reporting System)というシステムが設けられている。この3つのシステムの概要を述べる。

SAROAD<sup>15) ~ 19)</sup>は全国の大気汚染測定局で測定されたデータを蓄積する中央集中的なデータベースである。SAROADは3種類のデータ・ファイル (Raw Data File, Statistical Summaries File, Site Description File) から構成されている。Raw Data File には、連邦政府、州政府その他の機関における9000局 (常時測定局4000局) からの5000万の大気汚染測定データが蓄積されている。少なくとも3ヵ月に1度、各機関から測定データがEPAに提出され、所要のデータチェックを受けた後、データベースに蓄積される。データの量が膨大となるため磁気テープに入力、保管されており、そのためオフラインでの利用が可能である。一方、Statistical Summaries Fileには、raw dataのうち十分なデータ量があるものについて、平均、最大、その他の統計量を算出したものが収録されている。これはディスクに格納されていて、オンラインの利用が可能である。また、Site Description File には、測定点の位置の情報 (住所、緯度・経度等)、管理者、測定点の存在している都市の人口など、測定点に関する属性情報が収録されており、これもオンラインで利用することができる。

測定局コード、測定項目コードは、それぞれ9桁、5桁のコードで、全国的に統一したものが採用されている。コードの設定はデータセンターで行っている。SAROADは全国の大気汚染状況の把握、傾向の評価、各種規制の効果の評価、各種規制の効果の評価などのために利用されるほか、州や地方の関係機関において、データ集計やデータ解析に利用される。このため、これらに必要な各種の機能が内蔵されている。ユーザは各種のレポート型式の中からあるものを指定することにより、望む形でデータを取り出すことができる。データは、会話形式、リモートバッチ、郵送の3形式でユーザに提供される。

NEDS<sup>13) 14)</sup>は全国の大気汚染の発生源についての情報の蓄積、検索システムである。入力されている発生源の数は、100,000の点汚染源と3,200の

面汚染源（郡単位）である。それぞれの汚染源についてはばい煙発生施設の操業状態等に関する約 80 の項目が情報として入力されている。たとえば、個々の発生源の汚染物質の年間排出量は、これらの入力データから自動的に計算される。データファイルの作成は、各州から半年ごとに EPA Regional office を通して EPA に送られてくる（これは法律で義務づけられている）データをコンピュータでチェックし、チェック結果を各州に送り返し、訂正されたデータの返送を受けるといった操作を経て、行われる。

AEROS<sup>13) 14)</sup> はレポーティングの機能を有するシステムである。次のような目的に用いられる。①環境基準の維持達成のための施策の評価、②新規発生源に対する国の規制基準および自動車排出ガスの基準の強化にあたっての排出量および防止装置の評価、③ばい煙発生源の検査結果等の記録と蓄積

## (2) 水質関係（表 3）

水質汚濁に関する情報システムとしては表 3 に掲げるようなものがあるが、その代表的なものとして、STORET、PCS、WATSTORE と称されるデータベースがある。STORET (Storage and Retrieval for Water Quality Data) 全国の水質環境データ等に関するデータベース、PCS (Permit Compliance System) は水質汚濁発生源情報のデータベースで共に EPA により開発、運営されている。WATSTORE (Water Data Storage and Retrieval System) は内務省地質調査所により開発、運営されているもので、陸水、地下水の水質、水量等のデータに関するデータベースである。

STORET<sup>20) ~ 23)</sup> は全国のさまざまな機関で測定された水質環境データを納めるための、いわば「器」とであるとされている。STORET のデータセンターである EPA では、「器」を構成するコンピュータのハードウェア、ソフトウェアの充実と、ユーザに対する利用方法の普及や講習等を実施している。コンピュータそのものの管理は外部機関に委託している。そして、個々のデータの入力、維持管理等はすべてユーザの責任において、端末機を通して行われる。つまり、ユーザは自分のデータを STORET へ入力し、データチェックをして、データの修正、削除をする。したがって、データの正確さは、ユーザにすべて

依存することになる。このようにして STORET に蓄積されたデータはユーザ全員で共用され、それぞれの目的に利用される。STORET に蓄積されているデータファイルには、水質環境データファイル (Water Quality Data File) 以外に、家庭下水や工場排水の処理施設の情報についての Waste Facilities Inventory File、魚類の死滅事故に関する情報についての Fish Kill File、下水処理場の建設コスト等の情報についての Contract Awards File の 3 つの File がある。

水質環境データファイルは、米国の海洋、湖沼、河川等で測定された水質のデータと、その測定点の属性情報 (名称、位置など) を入力したものである。入力されている総データ数は約 4000 万件で、20 万測定点で 1800 項目にわたり測定されたデータである。データの入力は、データを測定した人 (機関) で STORET を利用したいと思う人 (機関) がそれぞれ行う。入力するデータは測定点属性情報と測定データである。測定点属性情報は次のものを入力する。測定者 (機関) を示すコード (これはデータセンターがそれぞれの測定者 (機関) ごとの個有のコードを設定する。)、データ入力者の名称、所在地、電話番号、アンロッキング・キー (公開したくないデータがあるとき、他の利用者に利用されることを防ぐためのデータ入力者自身がきめるパスワード)、測定点コード、州・郡・市コード、緯度・経度 (0.1 秒の精度で入力される)、リバー・マイル・インデックス (河口から測定点までの距離を表わすインデックス)。測定データは次のものを入力する。採水水深、測定項目コード (同じ項目でも測定法が異なるものは異なるコードとなる。)、測定データ (有効数字 4 桁以内)。データが入力された時点で測定項目ごとにデータチェック (上限値と下限値による) がなされる。ユーザが、STORET に蓄積されたデータを利用しようとする場合には、端末機により、まず行おうとする集計・解析の内容に合うプログラムを指定し、次いで測定点、測定データの指定をし、出力様式を指定することにより、望む型式で結果の出力を得ることができる。

PCS<sup>13) 14)</sup> は水質汚濁発生源情報のデータベースである。排水を放出する施設は連邦水質汚濁防止法に基づく許可が必要であるので、その施設許可の書類

からデータを取り出すことによりデータファイルが作成されている。そして機能としてレポーティング機能を備えている。現在、データ量として60,000施設分が入力されており、委託契約業者により毎週1回更新される。データはカードでバッチで入力される。出力形態として、施設に関する全情報(Facility Report)及び各施設についての情報を一行で出力したもの(Quick Look)などがある。

WATSTORE<sup>13) 14)</sup>は表流水、地下水の水質、水量に関するデータのデータベースである。蓄積されているデータは、測定局情報(測定局コード、測定場所等)、河川、湖沼の水量、水温等の日間平均値、流量の年間最大値、水質データ(200以上の水質項目についての測定結果のデータ)である。蓄積されているデータの量は、測定局に関するデータが10万以上、測定データが1億2000万である。

(3) 自然環境、資源、殺虫剤、その他関係(表4～表6)

自然環境、野生動物、資源、エネルギー、土地利用などに関しては、表4に掲げるようないくつかのデータベースがある。例えば、環境、経済、エネルギーの条件を入力すると環境政策が経済に与える影響をシミュレーションにより計算する機能をもつデータベースであるSEAS(Strategic Environmental System)やランドサット衛星写真情報のデータベースであるEROS(Earth Resources Observation System)などがある。殺虫剤等の化学物質に関する情報のデータベースは、表5に掲げるようなものがある。例えば、殺虫剤の毒性、使用状況についてのデータベースであるESPS(Epidemiological Studies Program System)や、殺虫剤の製造工場、製造量に関するデータベースであるERSS(Establishment Registration Support System)などがある。その他、法律の基準に関する情報、公害防止施設の予算に関する情報のデータベースもいくつかある。(表6参照)

(4) 文献情報、情報源情報関係(表7)

廃棄物、騒音、大気汚染に関してそれぞれ文献情報システムが設けられている。廃棄物に関してはSWIRS(Solid Waste Information Retrieval

System) というデータベース、騒音に関しては Noise File というデータベース、大気汚染に関しては APTIC (Air Pollution Technical Information Center) というデータベースがある。これらに入力されている情報は、文献題名、著者名、雑誌名などの書誌事項と要約である。APTIC の要約の部分は “Air Pollution Abstracts” として出版されている。水関係データの NAWDEX (National Water Data Exchange) というデータベースは、いわば情報源情報のデータベースといえることができる。すなわち、水関係のデータを有する機関は数多くあるので、どのようなデータがどこにあるかという情報をユーザに提供する。このデータベースには情報源ファイル (データを所有している機関の名称、タイプ、活動についての情報及びデータの提供機関についての情報、データの対象地域、データの種類についての情報) と、データのマスターファイル (測定点の位置、状況、データ測定機関、データの種類、データの利用可能な期間、測定項目、測定の頻度、データ蓄積形態、活動の目的、状況等に関する情報) の 2 つのデータファイルがある。

### 3.2 カナダ

カナダの環境情報システムとしては、WATDOC<sup>24)</sup> (Water Resources Document Reference Center) が代表的である。WATDOC は漁業環境省 (Federal Department of Fisheries and the Environment) が所管している情報システムで、1970 年に発足し、1972 年にオンラインサービスを開始している。現在では全国にユーザがいて広く利用されている。WATDOC のデータファイルには、CANADA WATER、D-REF、NEWS の 3 つがある。

CANADA WATER は関係の文献情報ファイルである。WATDOC のスタッフは、特定の雑誌、会議紀要および関係機関から送付されてくる文献、資料から水関係の文献を抽出し、分類、整理し、キーワード、コードの付加、要約の作成を行う。1970 年以降のカナダの水資源に係る文献 (出版されたものおよびされていないもの) が収録されており、現在このファイルは 27,000 レコードの規模である。



D-REF(Data-Reference) は、情報源の情報ファイルである。水資源関係のデータを保有している機関や人について登録している。情報の内容は、測定項目、方法、測定時間、利用の場合の制限、調査のスポンサー、所有者の名前、データの地理的範囲など26項目であり、これらの調査は、該当機関の担当者に直接面接することにより行っている。ユーザが情報を入手したい場合ユーザはこのファイルから所要の情報源を検索し、直接情報源に情報の提供を交渉する。

NEWSは、新聞の水関係のニュースを1行ずつの要約にしたデータファイルである。1970年5月からWATDOCのスタッフがカナダの新聞から水資源問題、水管理、その他水に関連したニュースを抽出して作成している。それぞれのニュースごとに9ケタの番号がつけられている。情報を入手したい人にはWATDOC 本部でコピーサービスが受けられる。

### 3.3 西 独

西独の環境情報システムとしてはUMPLIS<sup>25)</sup>をあげることができる。UMPLIS (Umwelt Planungs Information-system)は環境計画情報システムまたは環境情報ドクメンテーションシステムとも呼ばれる総合的な情報システムで、法律に基づいて作成、運営されている。UMPLIS はデータそのものを蓄積するデータバンクであると同時に、文献情報、情報源情報のデータバンクでもある。これは4つのサブシステムをもっており、各サブシステムの下にはいくつかのデータバンクがある。4つのサブシステムは次のようである。

- ① 環境全般に関するデータバンク
- ② 環境各分野に関するデータバンク
- ③ 環境に関する文献、法規、ドクメンテーション
- ④ 環境専門図書館

サブシステム④の下に存在するデータバンクとその運営状況は次のようである。「環境研究データバンク」においては、1975年以来行われて来た環境に関する調査研究に関するデータの蓄積が完了している。これは現在のところ、約

7,000の研究計画と約2,000の研究施設についての記述を含んでいる。データバンクに蓄積された情報は一定の様式に整理して、環境研究カタログ（略称UFOKAT）として出版されることとなっている。

「関係機関データバンク」に関しては、環境に関連のある官庁についての目録調査を行い、便覧としてまとめている。この目録は、環境保護の分野に関する仕事にたずさわっている約700の連邦および地方の官庁に関する情報（官庁の名称、住所、責任者名、法律的根拠、たずさわっている環境の分野、活動内容等）が示されている。

「環境モデルデータバンク」については、約800の数学的環境モデルを国内外から調査し、それぞれについて短い説明と書誌的情報（タイトル、モデルの製作者、機関、関係出版物）を収録している。

サブシステム⑥の下に存在するデータバンクとして、「廃棄物データバンク」、「水質データバンク」、「大気汚染データバンク」がある。たとえば、「廃棄物データバンク」は約1,000件の廃棄物の処分計画、利用計画、127件の廃棄物の種類、特徴、135件の廃棄物利用、147件の廃棄物利用技術についてデータが蓄積されている。

### 3.4 スウェーデン

スウェーデンにおいては、1970年から国家的に環境情報システムについて検討が開始された。そして、①化学物質、②発生源と汚染物質、③汚染地域および非汚染地域における動植物サンプル、④環境関係研究、についてのデータと情報システムの整備を行う方向が打ち出され、一部事業が実施されている。1974年からはSwedish Council of Environmental Informationにより環境情報システム（EIS）<sup>26)</sup>の整備が行われている。EISの中の代表的なサブシステムの概要は次のとおりである。

発生源データファイル：環境保護法に基づき規制対象となる汚染発生施設のデータが対象であり、手始めに、1975年～1977年にかけ3州の発生源（約1,000施設）についての約100,000の排出濃度データを収録した。今後全

国に拡大する計画であり、この場合データ量は約 10 倍となると見積られている。データのコンピュータ処理の内容として、日常的な解析、報告書の作成のほか、測定頻度、上限、下限などの自動チェックが行われる。

労働環境データファイル：労働環境に関するデータの収集、蓄積、処理、提供を目的とするもので、入力情報は、作業場に関する情報と、労働環境と健康の関係についての情報である。作業場に関する情報としては、会社の場所、従業員数、会社の健康管理、化学物質の使用状況、使用している装置、その他の化学的、物理的、機械的危険要因などが対象である。

環境研究インデックスファイル：環境科学の研究についての情報を蓄積、提供することを目的とするものである。対象は、一般の環境問題と労働安全衛生に関する研究プロジェクト、研究レポートである。約 4,000 件の環境研究（そのうち 2,000 がプロジェクト、2,000 がレポート）が登録されている。

#### 第 4 節 結 語

本章では、わが国および先進諸外国における環境情報システムに関し、その開発、運営の状況をレビューした。まず、わが国の環境情報システムの現状について論じ、文献情報、情報源情報については実用的に運営されているものが存在すること、しかし数値データについては今後の課題であることを述べた。ついで、先進諸外国における環境情報システムの実態を調べ、米国では、大気、水質、自然・資源・エネルギー、化学物質、法律・予算等についておよび文献情報、情報源情報について多くのデータベースが政府機関によって作成され、運営されていること、またカナダ、西独、スウェーデンにおいてもそれぞれ政府ベースで全国規模の環境情報システムが設けられていることを述べた。これらの中で、米国の大気環境データについてのデータベース、SAROAD、水質環境データについてのそれ、STORET は総合的なシステムで、全国的規模のものであり、今後のわが国の環境データベースの開発に関し有益な示唆を与えるものとして特筆した。

## 第4章 環境データベースの作成と環境問題への適用

### 第1節 概 説

全国で測定されている膨大な量の環境データが、組織され、コンピュータによる集中管理により、総合的に処理、利用されるようになることが期待される。前章で述べたように、米国をはじめ先進諸国では、それぞれのニーズに基づいて、環境データに関するデータベースが開発、運営されている。わが国においても、環境問題の研究や施策を実施する上で、このようなデータベースが必要と考えられる。そこで、本章では、環境データベースの意義について論じ、その有すべき機能について検討したのち、大気環境データと水質環境データを全国的規模で収集しデータファイルを作成することにより環境データベースを作成する。ついでこれによりデータがどのように効率的に処理、出力されるかを実例で示すとともに、環境データベースの今後の課題について述べる。

### 第2節 環境データベースの意義

#### 2.1 データベースの定義

鈴木<sup>1)</sup>の定義によると、データベースはさまざまな利用者がそれぞれの目的に応じて使用するデータを有機的に総合したものであるという。データベース操作システム (Data Base Management System, DBMS) は、この総合されたデータの操作に必要な機能をそなえたソフトウェアとハードウェアの組みであるとしている。データベースは DBMS がなければ機能せず、相互に一体のものとしてとらえ、両者をあわせたものをデータベースと呼ぶこともある。以後の議論においてはデータベースを有機的に統合されたデータの集合と DBMS の 2 つの部分からなるという考え方をとる。言いかえるとデータベースを作成ということは、有機的に統合されたデータの集合を作ることと DBMS を作ることになる。<sup>2)</sup>

前者は、大気環境データ、水質環境データ、廃棄物データ等の各環境データについて、その組織化を検討し、全国的に総合化され利用上便利なデータファイルを

作成することである。これについては次節で述べる。後者は、情報工学や数学の分野においてすでに多くの研究がされているところであり、本章の直接の検討対象としてはとりあげない。

## 2.2 環境施策とデータベースの関係

適切な環境施策の実施は関係情報の正確な把握と適切な解析・評価なしには不可能である。環境施策が高度化してくると、必要とする情報も多くなり、その処理方法も高度化したものが要求される。必要とされる情報は、数値データ、文献情報、情報源情報を含むものであるが、ここでは数値データを対象とし、環境施策と環境データがどのような相互関係を有するかについて大気汚染の場合を例として概念的に見てみる（図1参照）。

環境問題は、発生源からの汚染物により環境域が汚染され、環境域の汚染により、人間、動物、植物等が影響を受けるというパターンで現われる。したがって、環境問題の対策の基本の考え方は、「発生源」、「環境域の汚染」、「影響」の相互間における関係を解明し、「影響」が一定レベル以上にならないように「環境域の汚染」を抑制し、そのような「環境汚染」のレベルを維持するに必要な「発生源」対策を行うということとなる。このような前提において環境施策と環境データ関係の概念を図1にもとづいて述べると次のようになる。「発生源」、「環境域の汚染」、「影響」の相互間の関係の解明のためには影響発生メカニズム、汚染現象のメカニズムについて調べる。そして「環境域の汚染」のレベルを適切な評価指標で評価し、その結果、発生源に対して必要な行政上の措置（指導又は法律に基づく規制）を行う。これに対して発生源サイドは除去施設の設置、工程変更、燃料変更、発生源の再配置等の対策を考えることとなる。このような措置がとられた後の発生源における汚染質の排出と、これによりもたらされる環境汚染、そしてこれにより生じる環境影響について予測し評価する。評価が満足なものである場合はよし、そうでない場合は、前述のプロセスに適切な修正を加え再び繰り返すこととなる。

このような考え方に基づく環境施策の実施を支援するために環境データベース

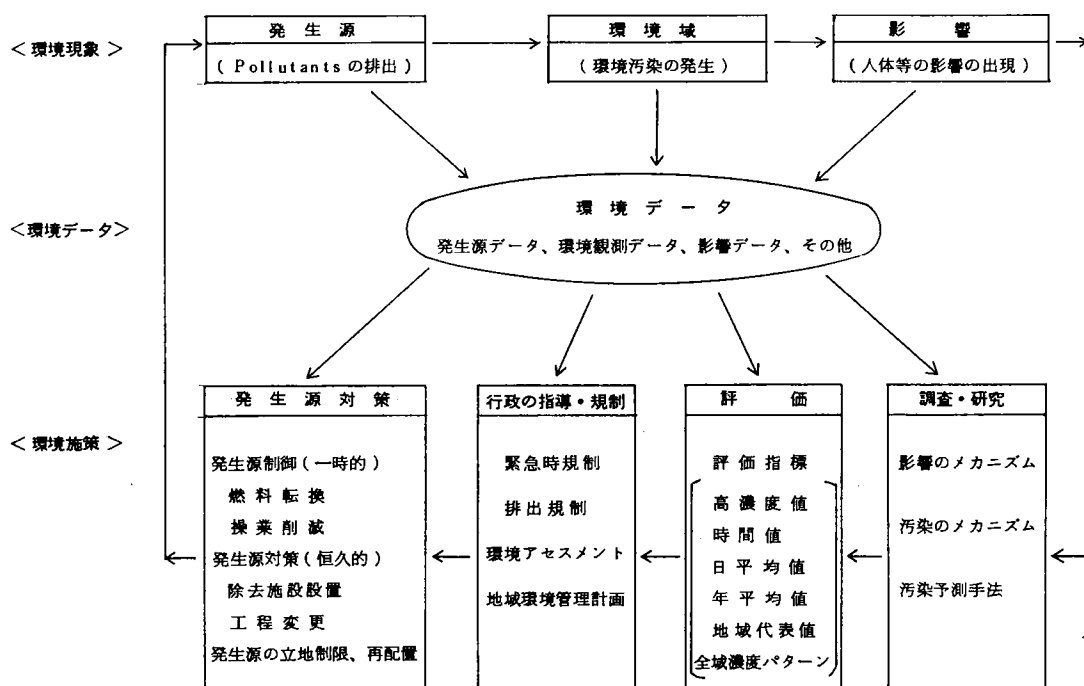


図1 環境データの環境施策に対する役割

は必要な機能——汚染質の排出、環境汚染の発生、環境影響の出現というような環境現象の状況を、発生源データ、環境観測データ、影響データ等を用いて、できるだけ正確に表現すると共に、正しく評価する。そして、相互関係の解析を通じて、必要な環境対策のための判断材料を提供するなどの機能——を有していることが前提である。

## 2.3 環境データベースの位置づけ

環境施策等の対象とするデータの範囲は非常に広い、たとえば大気汚染や水質汚濁に関する環境データのほか、自然環境データ、気象・水象データ、社会・経済関係データ、産業・工業関係データなどが対象となる。このような広い範囲のデータを一つのデータベースに蓄積することは困難であるし、また、仮にできたとしても既存のデータベースと重複することも考えられる。そこで関連のデータベース間のネットワークを考え、必要なデータは該当のデータベースから提供を

受けるというシステムを考えるのが实际的である。国家的な規模のデータベースの場合には、各省庁がそれぞれ所管事業に関するデータベースを作成し、これを互いにネットワークで結びつけ、データのやりとりをするのが望ましい。この場合、データのやりとりは、オンラインによるデータ交換のみならず、データファイルを磁気テープで交換する方法も実用性が高い。それぞれのデータベースの対象とするデータは所管のデータを中心にして一定の範囲のものに限られる。環境データベースの対象データは、大気汚染、水質汚濁等の環境観測データや発生源データ等のデータを主たる対象とする。しかしながら、このような種類のデータに限っても、全国のデータを時間的、空間的に網羅しようとするとな膨大なものとなる。そこで、中央レベルのデータベースと地方レベルのそれとの間の機能分担を考えるのが望ましい。次に、両者のデータベースの将来のあり方についての一案を述べる。中央レベルのデータベースの下に地方レベルのデータベースが位置するという構成を考える。しかし、それぞれ独立したものである。まず、蓄積するデータについては、両データベースの間でできるだけ同一のデータファイルの重複保有を避けるように考え、たとえば、地方レベルのデータベースの対象データは当該地域に関する過去から現在までのデータとする。中央レベルのデータベースの対象データは全国のデータとし、2次データ（平均値や最大値など集計された値）は全て、原データは特定の年度、期間（たとえば最近数年間）のデータとする。データは地方公共団体で測定され、そのデータベースに蓄積される（データベースがない場合においても何らかの形で保管、蓄積される）。データは毎年国に提出、報告され、このデータは中央のデータベースに蓄積される。ある地方公共団体が他の地方公共団体のデータまたは全国のデータの利用を望む場合、中央のデータベースからそれが提供される。データの処理、集計、解析のプログラムは中央で一括して作成し、これを各地方公共団体に提供する（こうすることによりデータ処理、集計の方法の全国的統一化が促進される）。したがって、中央と地方のデータベースの相互関係はデータは地方から中央へ、プログラムは中央から地方へという相互依存型の関係となる。

### 第3節 環境データベースの作成

#### 3.1 作成するデータベースの機能

データベースにどのような機能を設けておくべきかは、ユーザがこれをどのように利用したいと考えているかに依存する。ユーザには大別して2種類ある。まず、第一のユーザは、通常のプログラミング能力を有するユーザ（以下、プログラミング・ユーザと称する。）であり、他のユーザは、プログラミング技術を有していないが端末機を使って簡単なパラメータを指定するだけでデータ検索をしようとするユーザ（パラメータ・ユーザと称する。）である。環境データベースのユーザには、研究者、行政官のほか、コンサルタントや学生や一般人などが考えられるが、その利用ニーズからみると研究者グループはプログラミング・ユーザであり、行政官グループはパラメータ・ユーザであるといえることができる。この2種類のユーザがデータベースに対して望んでいる内容はかなり異なる。前者のデータベースに対する要望は、組織化されたデータの入手が容易になることやアプリケーション・プログラムが整備されることによりデータ解析が便利になることがあげられる。後者の要望は、全国規模の多量のデータの中から所定のデータが迅速に検索できることおよび多量のデータの編集、解析が迅速かつ正確にできることなどがあげられる。

以上の観点から、環境データベースの機能として、少なくともプログラミング・ユーザ用にはデータの処理、解析用のアプリケーション・プログラムが、パラメータ・ユーザ用にはデータの検索、処理、出力のためのソフトウェアが整備される必要があると考えられる。

プログラミング・ユーザ用のアプリケーション・プログラムとしては科学技術計算用のプログラムライブラリおよび環境問題の研究に用いられることが多い統計解析プログラムや汚染シミュレーションプログラム等があげられる。パラメータ・ユーザ用のソフトウェアを検討するには、行政上の業務で用いられるデータの検索、編集、解析の型式（以下、データ処理型式という。）を調べ、使用頻度の高いデータ処理型式について分類、パターン化する。このようなデータ処理型式についてはユーザは端末機からパラメータ指定を行うだけで所定の結果が得ら



れるシステムを考える。使用頻度の高いデータ処理型式は、ユーザのニーズ調査に基づいて定められる。大気環境データ、水質環境データのそれぞれについてユーザのニーズ調査を行い、特定のいくつかのデータ処理型式を設定した。このようなデータ処理型式を用いることにより、特殊な用件の場合を除けば、行政上のニーズの大部分に対応することが可能であると考えられる。

#### (1) 大気環境データの場合

大気環境データは、1年間、1測定局のデータでも非常にデータ量は多量であるので、これらの編集と統計量の計算がユーザの主要なニーズである。これを中心に、ユーザが一定の型式のデータの検索、編集、解析をするのが容易となるよう20種類のデータ処理型式（表1参照）を定め、このためのソフトウェアを作成した。用いるファイルは時間値ファイルで、指定するパラメータは、

- ① データ処理型式（データの検索、編集、解析のパターン）
- ② 測定場所（測定局あるいは都道府県、市町村）
- ③ 測定期間（年、年月、年月日、期間）
- ④ 測定項目

である。これを、データファイル検索条件式の形で書くと、次のようである<sup>3)</sup>。

$$a \triangle \left\{ \begin{array}{c} b \\ \\ b_1 \triangle b_2 \triangle \cdots \end{array} \right\} \triangle \left\{ \begin{array}{c} c \\ \\ c_1 \triangle c_2 \triangle \cdots \end{array} \right\} \triangle \left\{ \begin{array}{c} yy \\ yy / mm \\ yy / mm / dd \\ yy / mm / dd \triangle yy / mm / dd \end{array} \right\}$$

a：データ処理型式パラメータ

b：測定場所パラメータ（自治体コード、測定局コード等）

c：測定項目パラメータ

yy, mm, dd：測定時期パラメータ（年、年月、年月日、期間）

△：スペースを意味する。

#### (2) 水質環境データの場合

水質環境データの場合は、1測定点あたりの1年間のデータは大気環境データと比べて非常に少ない。しかし、測定点の数が非常に多いこと。水域単位

表 1 パラメータ・ユーザ用大気環境データのデータ処理型式

No.	編集出力様式名称	コード	処 理 内 容
1.	時間値月報リスト	A	1 か月分の時間値とその平均値等の集計値を出力する。
2.	日平均値など年報リスト	B	1 年分の日平均値などと、その集計値を出力する。
3.	拡張統計処理	C	任意の期間の時間値、日平均値等を使って、上記の A、B の処理よりくわしい統計量を出力する。
4.	日内変動	D	1 か月あるいは 1 年分の時間値の日内変動を出力する。
5.	週内変動	E	1 か月あるいは 1 年分の時間値の週内変動を出力する。
6.	風向・風速別平均濃度	F	1 か月あるいは 1 年分の時間値を使って、風向・風速別の汚染質の平均濃度等の集計値を出力する。
7.	風向指定濃度度数分布	XNNE XNEなど	指定された風向での、汚染質の度数分布を出力する。
8.	自治体別日報リスト	G	自治体内の全測定局の指定された年月日の 1 日分の時間値を出力する。
9.	自治体別月報リスト	H	自治体内の全測定局について、1 か月分の日平均値等を出力する。
10.	自治体別年報リスト	I	自治体内の全測定局について、1 年分の月平均値等の集計値を出力する。
11.	時間値気象月報リスト	J	風向、風速について、上記の A と同様に出力する。
12.	日平均値など気象年報リスト	K	風向、風速について、上記の B と同様に出力する。
13.	風向・風速別平均濃度あるいは風向度数分布	L	任意の期間における風向・風速別平均濃度を出力する。この出力様式は、出力様式 F と同じである。 ただし、風向を指定した場合には、風向・風速別の度数分布を出力する。
14.	自治体別気象日報リスト	M	風向、風速について、上記の G と同様に出力する。
15.	自治体別気象月報リスト	N	風向、風速について、上記の H と同様に出力する。
16.	自治体別気象年報リスト	O	風向、風速について、上記の I と同様に出力する。
17.	度数分布	P	任意の期間の度数分布を出力する。
18.	相関回帰分析	Y1 Y2 Y3 Y4	任意の期間の時間値等を使って、相関回帰分析を出力する。
19.	ベスト30、ワースト30	Z Z1 Z2 Z3 Z4 Z5	収録されている測定局について、時間値等でワースト30、ベスト30の測定局をリストする。
20.	測定局・測定項目一欄表	S	測定局の名称と測定項目の一欄表を出力する。

データの検索、集計が必要なことなどのため、このソフトウェアはデータそのものの検索、出力方法に主眼を置いたものとした。データ検索はハードウェアに付置されている DBMS を用い、出力の作図、作表のプログラムを新たに作成した。用いるファイルは水質データファイルと水質マスターファイル（次項参照）で、端末機からデータ検索の条件を指定することにより、望むデータが望む形に出力される。条件を指定するパラメータには、34 種類あるが主に次のようなものがある。

- |                     |   |   |
|---------------------|---|---|
| ①地域の指定              | { | <ul style="list-style-type: none"> <li>○ 都道府県名</li> <li>○ 水域名（河川、湖沼、海域の区別）</li> <li>○ 河川名、湖沼名、湾名</li> <li>○ 測定点コード、測定点名称</li> </ul> |
| ②環境基準に関する<br>条件の指定  | { | <ul style="list-style-type: none"> <li>○ 基準点、補助点の区別</li> <li>○ 環境基準類型</li> <li>○ 達成期間</li> <li>○ 環境基準指定年月日</li> </ul>               |
| ③位置に関する<br>条件の指定    | { | <ul style="list-style-type: none"> <li>○ 北緯（度分秒）</li> <li>○ 東経（度分秒）</li> </ul>  |
| ④データ集計に関する<br>条件の指定 | { | <ul style="list-style-type: none"> <li>○ 平均値、最大値、最小値等</li> </ul>  |

これらのパラメータを AND, OR の論理演算子を組みあわせて検索する。

### 3.2 データファイルの作成<sup>4) 5) 6)</sup>

大気環境データ、水質環境データについてはほぼ全国的規模と考えられるデータ量を収集し、第 1 章、第 2 章の検討内容に基づいて、実際にファイルを作成した。この概要を以下に述べる。

#### (1) 大気環境データファイル

大気環境データファイルとして、時間値ファイル、集計値ファイル、測定局ファイルと称する種類のデータファイルを作成した。時間値ファイルは原データで

ある 1 時間値データを収録したファイル、集計値ファイルは所定の統計量を計算した値を収録したファイル、測定局ファイルは測定局の位置やコード等の情報を収録したファイルである。これらの 3 種類のファイルを結びつけるものは測定局コードである。すなわち、測定局コードで検索し、3 つのファイルの内容を 1 つに編集し直せる。

#### ① 時間値ファイル

大気環境データの原データである 1 時間値を収録するものであるので膨大なデータ量となる。このファイルの作成にあたっては、主要な 10 自治体から磁気テープに記録されたデータ（記録方式は各自治体独自のもの）を収集し、データの記録方式を、第 1 章で検討した標準的方式に統一的に変換することにより行った。このための変換・編集プログラムの規模は約 9,000 ステップ（PL/1）である。磁気テープへの記録方式は第 1 章図 3 で示したものである。データ量は 1 年分のデータで約 83 メガバイト（1 レコード：17892 バイト、平均測定項目数：8、全測定局数 580）である。

#### ② 集計値ファイル

使用頻度の高い統計量を予め計算し二次データとして収録したファイルである。収録されている統計量は、SO<sub>2</sub>、NO<sub>2</sub> など 12 項目の 1 時間値の月間平均値や最大値、年間平均値や最大値などである。ファイルのデータ量は 1 年分で約 8 メガバイト（1 測定局の情報：1 レコード、1 レコード：1000 バイト、延べ測定局数：8000）である。

表2 大気環境データの測定局の属性情報

	内 容
測定局の名称・コードの情報	(1) 測定局の名称（カナ文字，ローマ字） (2) 測定局に関する国立公害研究所コード (3) 測定局に関する自治体測定局番号
測定局の位置に関する情報	(4) 測定局所在地・設置場所 (5) 用途地域コード (6) K値規制区域コード (7) 測定局の存在する場所の1/25,000地形図図幅名称 (8) 測定局の存在する場所の1/25,000地形図・地図番号 (9) 経緯度法による地域メッシュコード
測定局の保守管理状況に関する情報	(10) 測定局設置年・月・日 (11) 測定局面積 (12) 測定局構造コード (13) 給排水施設の有無 (14) 高圧ガス格納室の有無 (15) 空調施設の有無とその設置年月 (16) データ処理方法の種別，テレメータ伝送方式の種別，テレメーター使用開始年月 (17) 採気口高さと採気口の複数個の有無，集合管の有無 (18) 保守管理の常駐巡回の区別，巡回の担当者，巡回回数 (19) 保守管理担当者
測定局付近の発生源の状況の情報	(20) 固定発生源の名称，方位，距離 (21) 道路・交差点の名称，交通量，方位，距離
測定項目及び測定機種に関する情報	(22) 測定項目コード，測定機製造メーカー名，型式

### ③ 測定局ファイル

測定局の属性情報（測定局の名称、位置、付近の状況、測定項目など、表2参照）を収録したファイルである。測定局の属性情報は組織的に把握される制度となっていないので、全国の自治体にアンケートを配付して調査した。これに基づき測定局の位置を経緯度法による地域メッシュコードで把握するとともに、10桁の標準コード（第1章で提案した）を付加した。ファイル

のデータ量は約 1.5 メガバイト（1 測定局の情報：1 レコード、1 レコード：1 0 0 0 バイト、測定局数：約 1 5 0 0 ）である。

## (2) 水質環境データファイル

ファイルを 3 種類のファイルで構成させる考え方は大気環境データのそれと基本的には同じであるが、形の上では、測定点ファイルと集計値ファイルを合体して 1 つのファイル（水質マスターファイルと呼ぶ）とし、これと水質データファイル（原データのファイル）の 2 ファイルとした。

### ① 水質データファイル

水質測定データは、毎年 1 回各自治体から一定の様式の調表により国に報告される。報告されるデータは測定点ごとに個々の原データとその集計値である。水質データファイルは原データを入力したもので、通年調査と通日調査（第 2 章参照）のデータをそれぞれ別ファイルとした。磁気テープへのデータの記録は第 2 章で提案したフォーマットにより行った。ファイルに収録した情報は、測定点コード、採水時、水温、流量、採水位置、各測定データである。ファイルのデータ量は 1 年分で約 2.6 メガバイト（1 測定点 1 サンプルングのデータ：1 レコード、1 レコード：1 8 2 バイト、測定点数：約 6 0 0 0、1 測定点あたりのサンプルング回数：2 4 ）である。

### ② 水質マスターファイル

データ集計値（平均値、最大値、最小値、 $n/m$  など）と測定点の属性情報を収録したファイルである。測定点の属性情報は、測定局コード、水系・河川名称、測定点名称、位置の緯度・経度、環境基準に関する情報などである。この情報の整備は、5 万分の 1 の地形図に全国の測定点をプロットしその緯度・経度を読みとり、またそれぞれの測定点に測定点コードを与えるなどの作業によった。ファイルのデータ量は 1 年分で 1.4 メガバイト（1 測定点の情報：1 レコード、1 レコード：2 4 0 バイト、測定点数：約 6 0 0 0 ）である。

## 第4節 環境データベースの利用

前節で述べたデータベースの機能により、どのような出力結果が得られるかを実データを用いて見てみる。パラメータ・ユーザの利用例として、大気環境データ、水質環境データの集計業務への利用例を、プログラミング・ユーザの利用例として、データの2次元表示による汚染濃度分布の把握の例をとりあげる。

### 4.1 パラメータ・ユーザによる利用例

#### (1) 大気環境データの検索・出力例

大気環境データについて、ユーザが端末機からパラメータ指定を行うだけで望む検索・出力結果が得られる例を示す。とりあげた例は次のようである。

○測定局 ……東京都庁前測定局（測定局コード1301101010）

○測定項目 ……オキシダント（測定項目コード OX）

○測定年度 ……昭和49年度

○データ編集出力型式 ……日平均値などの年報リスト（1年分データの日平均値などとその集計値を一定の様式で出力する。）

（これに相当するデータ処理型式は表1に示すコードBである。）

ユーザは端末機のキーを、

B△1301101010 △ OX △49

と押すだけで該当のデータの処理結果を得ることができる。この出力結果の一部を図2に示した。

( B )		(ニサハイネンチ ナト* キンヌウ リスト)																				
1974 年	ソクタイ コウモク DX	( PPHM )				ジ*カンタイ 5ジ*~20ジ*				(サイコウチ)				ソクタイキヨク メイトチヨウタイ				コート* 1301101010				
月	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN	FEB	MAR	CNT	SUM	AVE	DEV	MAX	MIN				
1	4.0	4.0	4.0	6.0	10.0	2.0	6.0	4.0	4.0	2.0	9.0	5.0	12	60.0	5.0	2.3	10.0	2.0				
2	7.0	3.0	5.0	6.0	12.0	5.0	6.0	6.0	3.0	4.0	4.0	6.0	12	67.0	5.6	2.3	12.0	3.0				
3	6.0	4.0	4.0	5.0	10.0	6.0	4.0	2.0	5.0	5.0	6.0	6.0	12	63.0	5.2	1.8	10.0	2.0				
4	9.0	3.0	4.0	3.0	7.0	7.0	9.0	4.0	5.0	4.0	5.0	8.0	12	68.0	5.7	2.1	9.0	3.0				
5	5.0	6.0	5.0	4.0	9.0	10.0	6.0	3.0	12.0	4.0	6.0	5.0	12	79.0	6.6	2.7	12.0	3.0				
6	5.0	6.0	5.0	2.0	13.0	6.0	4.0	4.0	6.0	5.0	5.0	8.0	12	69.0	5.7	2.6	13.0	2.0				
7	4.0	6.0	8.0	3.0	6.0	7.0	4.0	3.0	6.0	12.0	10.0	3.0	12	78.0	6.5	3.8	16.0	3.0				
8	5.0	5.0	4.0	4.0	6.0	4.0	9.0	7.0	7.0	9.0	3.0	7.0	12	70.0	5.8	1.9	9.0	3.0				
9	4.0	5.0	5.0	4.0	4.0	5.0	5.0	9.0	3.0	8.0	2.0	4.0	7.0	12	61.0	5.1	1.9	9.0	2.0			
10	14.0	4.0	3.0	2.0	4.0	5.0	5.0	4.0	7.0	7.0	5.0	10.0	12	70.0	5.8	3.2	14.0	2.0				
11	15.0	4.0	4.0	3.0	3.0	5.0	3.0	5.0	5.0	8.0	4.0	8.0	12	67.0	5.6	3.3	15.0	3.0				
12	10.0	3.0	4.0	3.0	3.0	5.0	3.0	4.0	4.0	8.0	9.0	5.0	12	61.0	5.1	2.4	10.0	3.0				
13	9.0	3.0	4.0	3.0	4.0	5.0	2.0	3.0	6.0	7.0	4.0	6.0	12	56.0	4.7	1.9	9.0	2.0				
14	3.0	4.0	7.0	4.0	14.0	6.0	4.0	3.0	11.0	9.0	10.0	6.0	12	81.0	6.7	3.4	14.0	3.0				
15	4.0	10.0	8.0	5.0	4.0	4.0	5.0	4.0	4.0	12.0	3.0	3.0	12	66.0	5.5	2.8	12.0	3.0				
16	4.0	7.0	8.0	2.0	2.0	3.0	2.0	4.0	7.0	5.0	4.0	4.0	12	52.0	4.3	2.0	8.0	2.0				
17	9.0	7.0	7.0	4.0	-	4.0	2.0	3.0	7.0	4.0	3.0	5.0	11	55.0	5.0	2.1	9.0	2.0				
18	6.0	14.0	2.0	4.0	-	5.0	6.0	14.0	7.0	6.0	7.0	4.0	11	75.0	6.8	3.7	14.0	2.0				
19	4.0	8.0	3.0	3.0	5.0	5.0	5.0	7.0	2.0	5.0	4.0	5.0	12	56.0	4.7	1.6	8.0	2.0				
20	5.0	4.0	3.0	5.0	6.0	4.0	3.0	7.0	7.0	4.0	5.0	6.0	12	61.0	5.1	1.6	8.0	3.0				
21	3.0	6.0	3.0	1.0	6.0	5.0	2.0	10.0	6.0	-	9.0	6.0	12	61.0	5.1	2.6	10.0	1.0				
22	3.0	3.0	7.0	2.0	6.0	3.0	6.0	4.0	6.0	-	11.0	4.0	11	56.0	5.1	2.6	11.0	2.0				
23	6.0	6.0	3.0	3.0	4.0	5.0	6.0	4.0	6.0	3.0	6.0	4.0	12	56.0	4.7	1.2	6.0	3.0				
24	5.0	5.0	3.0	4.0	3.0	4.0	3.0	5.0	13.0	9.0	7.0	4.0	12	65.0	5.4	2.8	13.0	3.0				
25	4.0	5.0	6.0	2.0	2.0	7.0	4.0	4.0	11.0	7.0	11.0	5.0	12	68.0	5.7	2.8	11.0	2.0				
26	6.0	4.0	5.0	3.0	3.0	9.0	6.0	7.0	4.0	2.0	5.0	5.0	12	59.0	4.9	1.8	9.0	2.0				
27	9.0	5.0	3.0	8.0	5.0	7.0	3.0	5.0	7.0	5.0	8.0	7.0	12	72.0	6.0	1.9	9.0	3.0				
28	9.0	5.0	4.0	6.0	3.0	8.0	3.0	9.0	9.0	8.0	5.0	8.0	12	77.0	6.4	2.3	9.0	3.0				
29	4.0	6.0	6.0	4.0	3.0	6.0	4.0	12.0	3.0	6.0	-	7.0	11	61.0	5.5	2.4	12.0	3.0				
30	5.0	4.0	4.0	2.0	2.0	8.0	6.0	3.0	5.0	4.0	-	5.0	11	48.0	4.4	1.7	8.0	2.0				
31	-	6.0	-	3.0	2.0	-	7.0	-	6.0	6.0	-	7.0	7	37.0	5.3	1.8	7.0	2.0				
CNT	30	31	30	31	29	30	31	30	30	31	28	31	362									
SUM	186.0		141.0		164.0		146.0		193.0		170.0			1975.0								
		165.0		113.0		165.0		161.0		192.0		179.0										
AVE	6.2	5.3	4.7	3.6	5.7	5.5	4.7	5.4	6.4	6.2	6.1	5.8							5.5			
DEV	3.0	2.2	1.7	1.5	3.4	1.8	2.1	2.8	2.6	2.7	2.9	1.6							2.5			
MAX	15.0	14.0	8.0	8.0	14.0	10.0	9.0	14.0	13.0	12.0	16.0	10.0							16.0			
MIN	3.0	3.0	2.0	1.0	2.0	2.0	2.0	2.0	2.0	2.0	3.0	3.0							1.0			

図2 データベースによる大気環境データの出力例

## (2) 水質環境データの検索・出力結果

次に、水質環境データの検索・出力例を示す。とり上げた例は次のようである。

○対象測定点 ……東京都内の荒川において環境基準類型がDに指定されている水域にある測定点

○データ編集・出力型式 ……全てのデータとその棒グラフにしたものを出力する。

データ検索条件式は

KEN="TOKYO"&KAWA="アラカワ"&RUIKEI="D"

となる。ユーザは端末機からこのようにパラメータを指定すると望む結果が出力される。出力結果の一部を図3に示した。



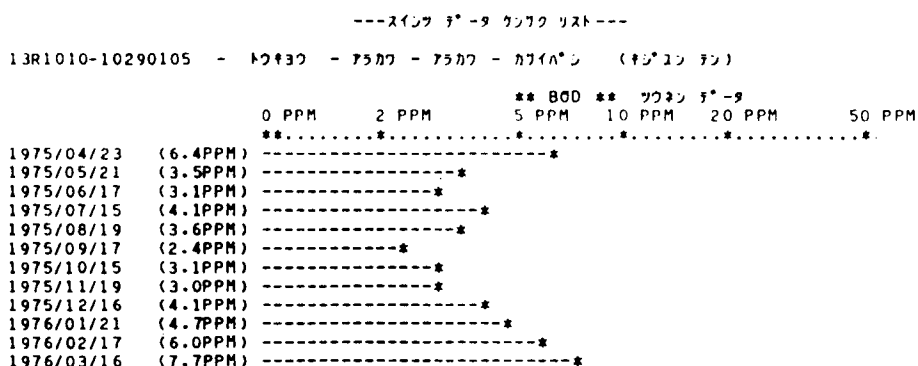


図3 データベースによる水質環境データの出力例

## 4.2 プログラミング・ユーザによる利用例

プログラミング・ユーザ用のアプリケーション・プログラムのうち一般科学計算用のプログラムの多くはハードウェアに付備されているプログラム・ライブラリの中に見い出さる。しかし、環境データに特有の解析技術のプログラムは環境データベースの機能の一部として独自に作成しておくことが望ましい。ここでは、データのメッシュデータ化（2次元表示）のプログラムの例を述べる。すなわち、離散的な数の限られた地点で測定された光化学オキシダントの汚染濃度データを面情報（2次元分布パターン）として表現する方法をとりあげる。データの2次元化はスプライン法によるデータ補間法による方法（第6章で詳細に論じる。）を用いる。これによって作成した時間ごとの分布パターンの出力図を図4に示す。これの特徴は汚染濃度の分布パターンが時間の経過に伴って変化してゆく状況が視覚的に把握できることである。とりあげたケースは関東平野（東京湾岸地域）におけるオキシダントの発生と推移の状況である。出力図から、「川崎市の海岸地域の一部で発生したオキシダントが正午頃にはかなり広い面積で高濃度となる。これは時間の経過とともに東京都世田谷区、横浜市港区方面へ広がり、スモッグの中心部も北へ移動する。午後4～5時には北上しつつ消滅する。」という現象状況が視覚的に理解できる。

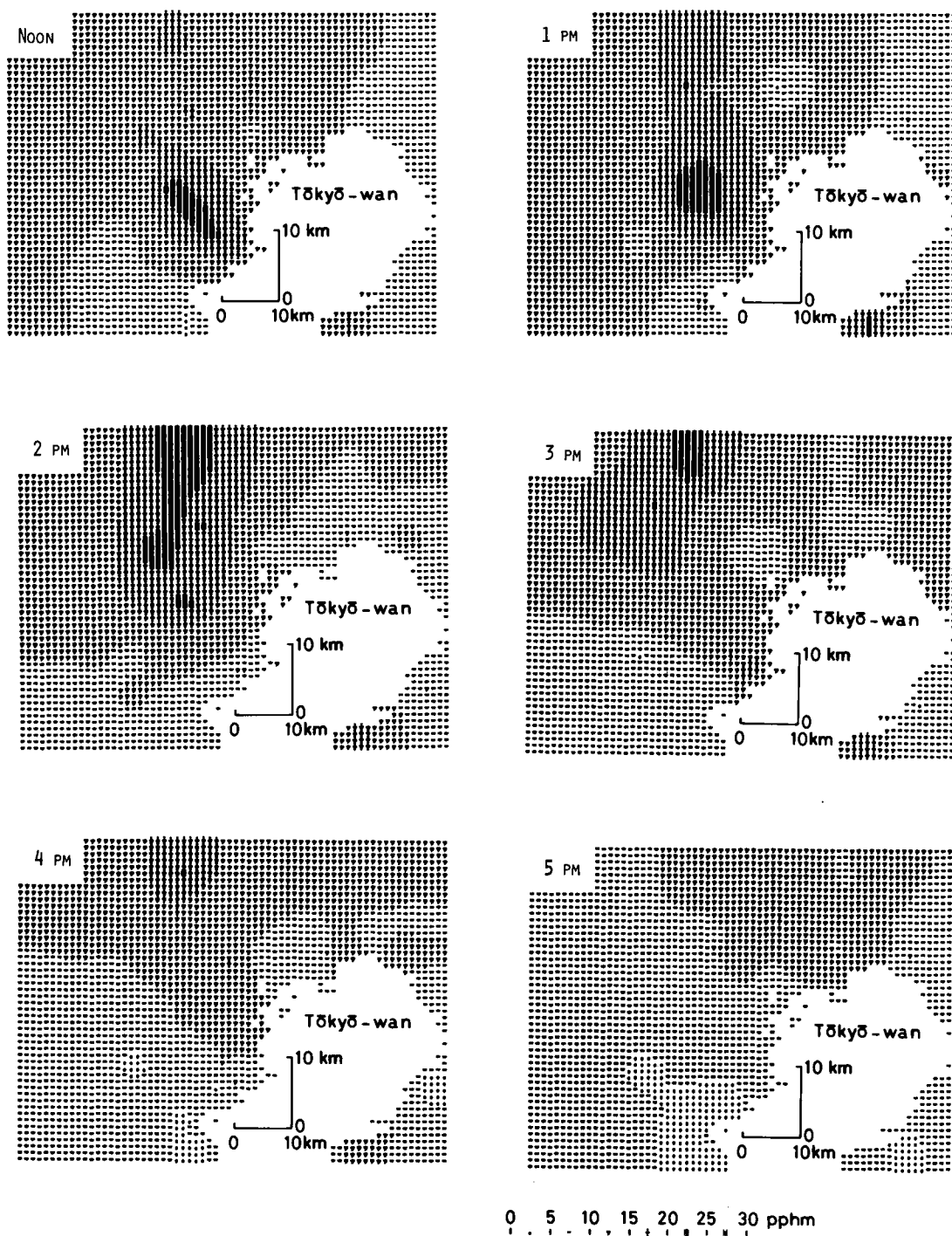


図4 データベースによるデータの2次元分布パターン表示出力例

(オキシダント濃度分布の経時変化)

## 第5節 データベースの環境問題への適用上の課題

環境データベースが環境問題の研究や施策の実施のためにより有効な武器になるためには、データ検索、処理、解析の機能が拡充されることが必要である。すなわち、次の事項について検討されねばならない。

まず、第1番目に、データ検索のためのソフトウェアの拡充である。ユーザが端末機からパラメータ指定を行うことにより望むデータの検索と望む形の出力結果が得られる方式の機能向上を図るためには、データ検索ソフトウェアの性能の改良を検討することが必要である。また、オンライン・データ検索の対象データはコンピュータの外部記憶装置（ディスク・メモリ）に蓄積されていなければならないが、データ量の増加が不可避である見通しから、その容量の拡大が検討されねばならない。

第2番目に、データ処理、解析のための機能の拡充、整備をはかること。最近ではコンピュータ・メーカーがユーザ・サービスの一環としてデータ解析用の数多くのプログラムをユーザの使用に供しているが、環境データの処理、解析のためのアプリケーション・プログラムは十分でない。このようなプログラムをできるだけ多く開発・作成することが望まれる。

第3番目に、データファイルの種類と量を増やすこと、たとえば、データ解析の目的によって、発生源データ、気象データ、水象データ、公害被害データ、その他（人口等の社会関係データ、工業統計等の産業関係データ、自然環境データ等）についてもデータファイルが必要な場合がある。これらのデータのデータベースシステムへのとり込みについて検討することが必要である。

以上の課題は、コンピュータ・メーカー、行政機関（データ・センターを含む）、環境科学者（ユーザを含む）等がそれぞれ努力することによって解決されねばならない。1番目の課題は主としてコンピュータ・メーカー側の課題であり、3番目の課題は主としてデータ・センター側の課題であるということができるが、2番目の課題は、とくに環境データの特性に起因して生じるデータ処理、解析上の問題の技術的対応等は、関係者において検討されるべきものである。この対象となるものは数多く考えられるが、実務上ニーズの高い次のような機能の

開発が例としてあげられる。

- ① 自動的にデータをチェックし、正しいものとそうでないものを区別し、修正する機能
- ② 離散的な地点で得られたデータ（点データ）をメッシュ点ごとに与えられたデータ（面データ）として表現する機能

## 第 6 節 結 語

組織化されたデータにもとづきデータベースを作成することにより、データの処理、利用がどのように便利になるかを論じた。まず、データベースの定義、環境施策とデータベースの関係、環境データベースの位置づけについて論じ、ついで作成するデータベースの機能及びデータファイルの内容について述べた。データファイルは第 1 章及び第 2 章で検討した磁気テープの記録方式に基づき大気環境データと水質環境データについて作成した。そして、作成したデータベースによりデータの処理や解析が便利に効率的に行えるようになることを実例（端末機からの指示によりデータを所定の内容で処理し所定の様式で出力する例や汚染度を 2 次元分布パターンとして表現する例）で示した。また、環境問題へ適用する場合のデータベースの機能の向上について今後の課題を 2 ～ 3 指摘した。

## 第Ⅱ編 環境データの環境問題への適用技術 に関する研究

## 第5章 データの自動チェックシステム に関する研究<sup>1) 2)</sup>

### 第1節 概 要

既に第1章、第2章において、データの組織化のための課題として、データチェックに関する問題の重要性を指摘した。測定データには各種の原因で異常値や誤差が混入することは避けられないので、適切なデータチェック作業を実施することによりこれらを取除くことが必要である。しかし、全国の関係機関の現状をみると、データチェックの作業量が膨大であること、チェックの方法が困難であることなどにより、十分適切なデータチェックが行われているとは言いがたい。なかにはコンピュータを用いたデータチェック方法を採用している機関もあるが、<sup>3)4)</sup>全ての機関で適切に行われることが必要であり、また全国的に統一的な方法で行われることが望ましい。そこで、本章では大気環境データを対象としてコンピュータを用いたデータの自動チェック方法（以下、Automatic Data Checking System または ADC と呼ぶ）による異常値の検出を検討する。ADC の採用は、膨大なデータのチェック作業から人を解放できること、経験と勘にもとづくチェックから機械的統一的なチェックへ変換することによりデータの精度を均質化することができることなどにおいて意義がある。

### 第2節 環境データの自動チェックの課題

#### 2.1 現行システムの問題点

大気環境データのチェックについての現行のシステムの問題点について述べる。各自治体で行なわれている大気環境データの測定からチェックが完了したデータの作成までのデータ処理過程は一般には次のようである。<sup>5)</sup>各測定局で自動測定器によって測定されたデータは、チャートに記録がとられる一方、テレメータシステムを通して、毎時間中央監視局に送られ、テレメータ制御用コンピュータにより磁気テープに記録される。（この記録内容を磁気テープAと呼ぶこととする。）ところで、このデータは測定機やテレメータ装置の状態、その他の原因により、

さまざまな異常値を含んでいるので、最終的に確定したデータを作成するには、データが1カ月程度まとまった段階で、データのチェックを行ない、異常なデータを検出し、修正する。この修正にあたっての情報は、測定機のチャートのデータ及び測定の保守・点検時の記録によっている。すなわち、自治体ではおのこの保守管理体制を有しており、通常、週に1回程度の割合で測定局の巡回を行ない、反応液、パイプ類、電気系統、その他測定機の状態およびテレメータシステムの状態を点検、調整しているが、この保守・点検時の記録を参考にして、異常値を検出し修正し、最終的に確定した1時間値データは磁気テープに保存用データとして記録される。以下、この記録内容を磁気テープBと称することとする。このデータ処理の作業は、より精度の高いデータを得る上で不可欠なことであるが、実際上は各種の問題点がある。主なものをあげると、

- ① 多くの人手と時間がかかる。
- ② 異常値が発生する原因（測定機の保守管理状況等）の具体的検討が必要となる。
- ③ 異常値とするかどうかの判断基準とデータを修正する方法（データ処理基準）が標準化されていない。

これらの問題点についてももう少し説明を加えると、データ処理の作業は、磁気テープAに入力されているデータと測定機のチャートのデータを一つ一つ突合することにより異常値を見つけ出し、修正情報（欠測も含む）をカードや紙テープで与えてデータを修正する。一般に一つの自治体で1カ月に1～2万（10～30測定局×約10項目×720時間）ものデータがありこれに対して上述のような作業をするには多くの人手と時間がかかる。修正情報を知るためには、異常値の発生原因の究明に立ち入らねばならない。異常値の発生原因は大きく分けて、測定機によるもの、テレメータシステムによるもの、外部からの影響によるものが考えられる。表1にそれぞれの場合の異常値発生原因の例を掲げた。これらの発生原因は常に明白な状態で現われるわけではない。異常値発生原因は多種多様であり、その程度（例えば流量変動、電圧変動の程度）も大きなものから小さいものまでである。また、場合によっては複数の異常値発生原因が重なることもある。どのような場合に、データを異常値あつかいにするか

という判断基準及び異常値を修正して正しい値にする場合の修正方法については、標準的な基準がなく、自治体によってやり方が異なったり、また担当者の個人の判断によっている場合も多い。チェックの程度の精粗によりデータの信頼性は異なるので、このようなデータにより全国的な汚染度の評価を行なうと評価自体

が不正確になる。全国的に統一されたデータ処理の基準に基づいてデータ処理を行なうことによりデータの均一性、信頼性を高めることが必要性である。

表1 異常値の発生原因の例

測定機	<ul style="list-style-type: none"> <li>・サンフル流量の変動</li> <li>・ゼロ点のずれ</li> <li>・リセット不良</li> <li>・反応液、酸化液等の液もれ、蒸発</li> <li>・表記部分の異常</li> <li>・電気系統の故障</li> </ul>
テレメータ	<ul style="list-style-type: none"> <li>・制御装置の故障</li> <li>・電圧変動による送受信時の誤差</li> </ul>
外部からの影響	<ul style="list-style-type: none"> <li>・妨害物質による誤差</li> <li>・近接煙源による影響</li> <li>・その他(落雷等の事故による影響)</li> </ul>

## 2.2 自動チェックの環境データへの適用の方向

コンピュータを用いた異常値の自動検出法の採用は、データ処理に関する前述のような諸問題をかなりの程度解決するのに役立つものである。ADC (Automatic Data Checking System) に大別して次の2方式がある。すなわち、コンピュータを用いて自動的に行なうのは疑問値(データチェック手法により異常値として検出されたものは疑わしいけれども必ずしも異常値とは限らない。これを「疑問値」という。)だけで、疑問値の検討、訂正は人手で行なう方式(部分ADCと呼ぶ)とADCで発見された疑問値はそのまま異常値としてとりあつかい修正値も自動的に与える方式(完全ADCと呼ぶ)である。大気環境データのデータチェックの場合、前者が採用されている。気象庁が行う気象データの処理では後者も採用されている。気象データの処理に完全ADCが採用されているのには理由がある。すなわち、雨量のような気象データを考えると、異常値として検出されたデータを全て観測値から除いてしまうより、たとえそれがラフな推定値であったとしても、異常値として検出されたデータの代りにその推定値を用いた方が統計量(例えば年間降雨量、年平均降雨量など)の精度は高まるからである。環境データの場合は、年平均値なども重要であるが、個々のデータが一定



値（環境基準や緊急時基準など）を超えるかどうかということの方が重要視されるので、その採用には十分注意が必要である。すなわち、この方式は、実際に測定されたデータを統計学的に疑問な値だからと言って、より正しそうな値に統計学的に変更する方法であるので、データの変更がデータのねつ造というような誤解をされることがあってはいけなからである。

このような観点から、環境データのチェックを自動化する方法を導入する当面の方向として、異常値の可能性のあるもの（疑問値）をコンピュータにより検出し、その疑問値について人間が真の異常値かどうかを判断する、そして異常値と判定したデータは欠測値扱いとするという部分 ADC の採用が、現時点で、最も社会的に受け入れられやすいものと考えられる。これが実用化されることの効果は、データ処理に関する入手の作業の大幅な削減とデータ処理の全国的標準化が促進されることである。

なお、データチェックは、その実施の時点からみると次のような 2 つの種類に分けられる。各測定局で測定されたデータが中央局に伝送されて来た段階で即時に行うものと、データの測定後ある程度の時間（例えば数週間から数か月）が経過した後行うものである。前者のように刻々のデータを即時にオンラインでリアルタイムにチェックする方式をリアルタイム・チェック、後者のようにデータがある程度たまってから一括してオフラインで処理するのをオフライン・チェックと呼ぶ。わが国における大気環境データのチェックでは後者は実施されているが、前者は実施されていない。環境データについて、後者（オフライン・チェック）の有効な手法を確立すること及び前者（リアルタイム・チェック）の可能性を検討することが重要な課題である。

### 第 3 節 データチェック手法の有効性の検討

#### 3.1 データチェック手法の種類と検討対象

観測データ群から異常値またはその疑いのあるデータを取り出す（異常値を検出する）方法は、対象データごとにその比較値を設定し、両者の差が一定以上である場合に、これを異常値（または異常値の疑いのあるデータ）とするという

ものである。データチェック手法は、比較値の設定方法によって次のように分類できる。<sup>1)</sup>

(a) 当該観測値がないものとして時系列データの相互の関係からその値を予測し、それを比較値とする方法（以下、これを時系列値チェック法と呼ぶ）。予測の方法により次のような方法がある。

① 過去の観測値の変動パターンの傾向や特徴を把握することにより当該観測値を予測する方法。予測手法として非物理モデルを用いる。

② 過去と未来のいくつかずつの観測値（既知）を加重平均することにより、当該観測値を予測する方法（以下、これに基づくチェック法を加重平均法と呼ぶ）。

③ 前時刻の観測値を当該時刻の予測値とする方法、これは言いかえると相隣合わせた時刻の観測値を比較する方法（以下、これに基づくチェック法を偏差法と呼ぶ）。

(b) 物理モデルにより汚染予測をし、これを比較値とする方法。大気拡散式などを用いて予測する方法が考えられる。

(c) 他のデータとの相関関係に基づき予測し、これを比較値とする方法（以下、これを相関予測チェック法と呼ぶ）。

① 同時に観測された他の測定項目の観測値との間の回帰により予測する方法。これに基づくチェック法を項目間回帰法と呼ぶ。

② 同時に他の測定局で観測された同じ測定項目の観測値との間の回帰により予測する方法。これに基づくチェック法を局間回帰法と呼ぶ。

(d) ある固定した値を比較値とする方法。これに基づくチェック法は、観測値が一定値を超えるかまたは一定値に満たない場合、その観測値を疑問値とする方法である。これを上下限值法と呼ぶ。

上記の方法のなかで実務上最も広く用いられているのは(d)であり、(a)の③もかなり用いられているが、その他は現時点では余り用いられていない。

以上述べたデータチェック手法のうち、本章で検討対象としてとりあげるのは、(a)時系列値チェック法、(c)相関予測チェック法、(d)上下限值法である。(b)物理モ

デルによる汚染予測に基づくチェック法は次の理由によりとりあげない。すなわち、物理モデルによる予測値の誤差は観測値自身に生じる誤差より大きい可能性が大であること、及び、汚染予測するのに発生源データや気象データなど多くの関係データを把握することが必要であることなどにより、データチェック法としては実用的でない。これに対して時系列値チェック法は、当該時系列の観測データ群のみを対象としてデータ予測が可能であり実用性が高いと考えられる。時系列値チェック法としては①、②、③の3種類の方法を、相関予測チェック法としては項目間回帰法を検討対象としてとりあげる。

### 3.2 有効性の評価尺度

データチェック手法の有効性をどのように評価すればよいかを検討する。

データチェック手法により検出された疑問値の集団の中には異常値が含まれているが、また、疑問値として検出されなかった測定値の中にも異常値がある可能性がある。異常値を検出する手法として好ましいものは、その手法を用いることにより、できるだけ多く（できれば全部）の真の異常値が検出でき、かつ余分な疑問値はできるだけ検出しない手法であるということができる。手法の有効性を評価するための判定基準として、二つの量を定義することとする（図1参照）。<sup>1)</sup>

a) 発見率：当該手法により、総数  $N_o$  個の異常値のうち  $N$  個が検出されたとき  $N / N_o$  を発見率と呼ぶ。

b) 適中率：当該手法で検出した疑問値（その数  $M$  個）のうち、真に異常値であった値を  $N$  個とすると、 $N / M$  を適中率と呼ぶ。

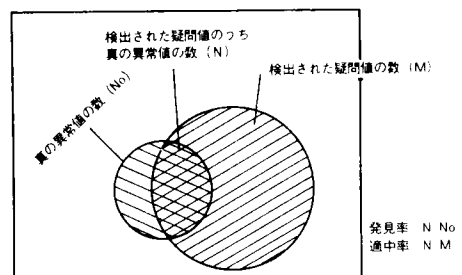


図1 異常値検出の発見率と適中率

これらの二つの量は共に0～1.0の間の値をとる。望ましいチェックの方式としては、発見率、適中率ともに高い（1に近い）ことである。両方とも1に近づけられれば理想的であるが、両者の関係は一方を高めようとするとは他方が低くなるという関係にある。すなわち図1で見られるように、疑問値検出率（疑問値

として検出したデータ数／全データ数)を大きくとれば、発見率は大きくなるが逆に適中率は小さくなる。疑問値検出率を小さくとれば、その逆である。したがって、この関係をグラフに描けば、図 2 のようになり、グラフの上方に位置する手法がよい手法である(手法 A は手法 B よりもよい手法である)と言える。しかし、手法の評価のために、常にこのように疑問値検出率を幾通りもかえて発見率、適中率を調べてみる必要は必ずしもない。ある疑問値検出率における発見率、適中率の両方の値を調べることにより、手法の有効性を判断することができる。

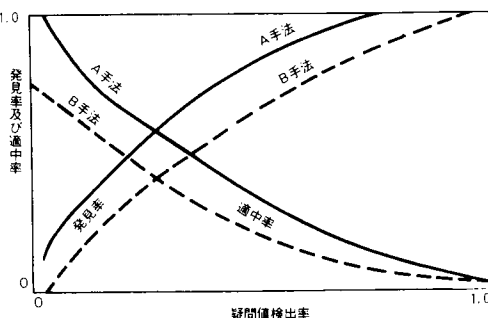


図2 疑問値検出率と発見率、適中率の一般的関係

### 3.3 ADF の適用の検討

時系列値チェック法のうち、過去の観測値の変動パターンの傾向や特徴を把握することにより次のデータを予測する方法として、ADF (Adaptive Digital Filter) をとりあげる。ADF は、過去のデータを学習させることにより、その傾向、特徴を把握させ、次のデータを予測するもので、もともと装置やシステムの故障診断などに用いられていた手法である<sup>6)</sup>。これ以外に採用しうる方法としては、ウィナーフィルタ (Wiener Filter)、カルマンフィルタ (Kalman Filter)、GMDH (Group Method of Data Handling) 等があるが、ADF を採用したのは次の理由による。ADF はシステム雑音や観測雑音に関する事前情報やシステムの構造に関する知識をほとんど必要としない。すなわち、データに関する相関関数、パワースペクトル密度その他システムに関する事前情報をほとんど必要とせず、またフィルタの構造が簡単であることである<sup>6)</sup>。また、この方法はデータの予測に未来のデータを用いる必要がないので、オフライン・チェックのみならずリアルタイム・チェックにも利用できる。

#### (1) ADF の原理<sup>6)</sup>

測定値を時刻順に  $x_1, x_2, \dots, x_N$  とするとき  $x_j$  の推定値  $\hat{x}_j$  が(1)式、(2)

式に示すように、その前の  $p$  時間の実測値  $x_{j-k}$  ( $k = 1 \sim p$ ) の 1 次結合、あるいは実測値  $x_{j-k}$  と推定値  $\hat{x}_{j-k}$  の 1 次結合で表わされるようなモデルを考える。前者を non-recursive モデル、後者を recursive モデルと呼ぶ。

$$\hat{x}_j = \sum_{k=1}^p a_k x_{j-k} \dots\dots\dots(1)$$

$$\hat{x}_j = \sum_{k=1}^p a_k x_{j-k} + \sum_{k=1}^p b_k \hat{x}_{j-k} \dots\dots\dots(2)$$

係数  $a_k$  ,  $b_k$  を推定値と実測値の差の 2 乗の期待値  $E(e_j^2) = E[(x - \hat{x}_j)^2]$  を最小にするという条件から決定するのが理論的であるが、この条件から直接係数の値を求めるかわりに、(3)式のように  $E(e_j^2)$  を用いて係数を逐次変化させてゆき、正しい解に近づけようという方法がある。しかし、 $E(e_j^2)$  を求めるのは困難なことなので、ADF では、簡便な手法として(4)式のように個々の誤差によって係数を逐次変化させてゆき、正しい解に近づけてゆくのである。

$$\left. \begin{aligned} a_{k,j+1} &= a_{k,j} - \hat{v} \frac{\partial}{\partial a_{k,j}} E(e_j^2) \\ b_{k,j+1} &= b_{k,j} - \hat{u} \frac{\partial}{\partial b_{k,j}} E(e_j^2) \end{aligned} \right\} \dots\dots\dots(3)$$

$$\left. \begin{aligned} a_{k,j+1} &= a_{k,j} - \hat{v} \frac{\partial}{\partial a_{k,j}} e_j^2 \\ &= a_{k,j} - v e_j \frac{\partial}{\partial a_{k,j}} e_j \\ b_{k,j+1} &= b_{k,j} - u e_j \frac{\partial}{\partial b_{k,j}} e_j \end{aligned} \right\} \dots\dots\dots(4)$$

ここで、 $k = 1. 2. \dots\dots p$  ,  $v = 2 \hat{v}$  ,  $u = 2 \hat{u}$  ( $u, v$  はきざみ幅で十分に小さい実数)

(4)式にもとづいてコンピュータにより繰り返し計算を行ない、収れんする値で各係数を定め、この係数を用いて式(1)又は(2)から推定値  $\hat{x}_j$  を計算する。この計算を時間ごとに行なう。

## (2) 実データへの適用

対象とするデータは、データチェックを施す前のもの、すなわち測定機で測定され、テレメータシステムで中央局に送信され、磁気テープに記録されたいわゆる未修正データ（前述の磁気テープ A）である。用いたデータは、神奈川県で測定された昭和 51 年度のオキシダントのデータである。

(1)の方法を用いて予測値を計算した。図 3 は観測値と ADF による予測値を並べて図示したものである。<sup>7)</sup>

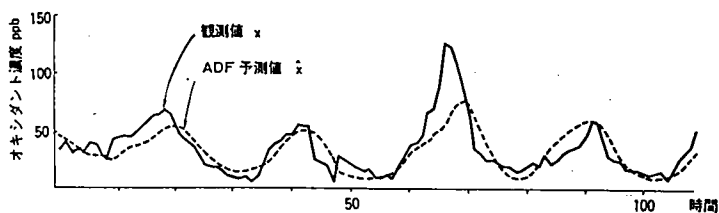


図 3 オキシダント観測値と ADF による予測値

この図で見られるように両者はよく一致しており、ADF が観測値によく追隨していることがわかる。ところによって追隨しえないところがあるが、こういうところが異常値が発生している箇所である疑いがある。両者の差 ( $e_j = x_j - \hat{x}_j$ ) をとって表示すると、ADF が追隨しえている通常のデータの箇所は差が小さく、図上では小さいふれ幅であるが、ADF が追隨しえなくなる部分は差が大きく、図上ではふれ幅が大きくなる（図 4 参照）。

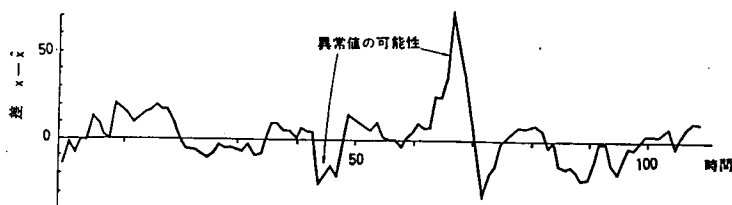


図 4 ADF によるシステム異常の検出の信号

このような場合、システムの異常（測定機の故障、電送装置の故障など）の発生が疑われる。どの程度以上のふれ幅のときに疑問値（異常値の可能性あり）とするかは、疑問値検出率（全データ数に占める検出される疑問値の数の割合）

のとり方の問題になる。疑問値検出率を余り大きくし過ぎでは正しいデータまで異常値として検出する可能性が増えるし、余り小さくし過ぎるとシステムの真の異常も看過してしまうことになる。したがって、目的に応じて適当な疑問値検出率を設定する必要がある。この問題は後に論じる。

次に、疑問値検出率と異常値の発見率、適中率の関係を調べる<sup>7)</sup>。チェック手法により異常値を検出したものが真に異常値かどうかの判断は真のデータを知らねばできないが、真のデータというのは知ることが不可能であるので、2.1で述べたような方法で人手によりチェックされたデータ（磁気テープB）を真のデータと見なし、磁気テープAの磁気テープBと異なる部分のデータを異常値とした。いくとおりかの疑問値検出率に対する発見率、適中率を求め図示したのが図5である<sup>7)</sup>。図から、3.2で述べたように、疑問値検出率を増加させると発見率は高まり、一方適中率は下がること、疑問値検出率を小さくするとその逆であることがわかる。疑問値検出率を0.02～0.04程度にとると、発見率、適中率は0.3～0.4程度、疑問値検出率を0.1程度にとると発見率は0.6～0.7程度になる。実際の発見率は次のような

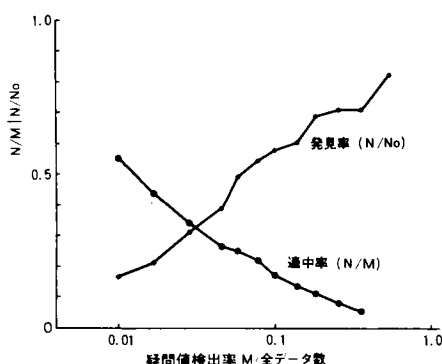


図5 ADFを用いた場合の疑問値検出率と発見率、適中率との関係

理由によりこれより高くなるので、この方法は十分実用の可能性がある。すなわち、異常値のなかには測定機の保守管理上発生する測定値誤差たとえば反応液の漏れや蒸発によって発生する測定値誤差——このような測定値誤差はそれが小さな場合はADCでは検出困難である——が含まれているが、このような異常値は発生原因が既知であり測定値の補正が可能であるので、オフライン・チェックの場合には、この作業がADCの実施とあわせて行われることにより発見率は高められる。

### (3) 疑問値検出率の設定

疑問値検出率をどの程度の値に設定するのが適切であるかについては、発見率にウェイトを置くか、適中率にウェイトを置くかによってその値は異なる。

たとえば、ADC によって検出された疑問値を人手でチェックする部分 ADC の場合は、できるだけ異常値を見落さないようにするため、発見率にウェイトが置かれ、疑問値検出率を比較的大きくすることとなる。その結果、適中率は低下するが、検出された疑問値を再度人手によってチェックすることによりカバーするという事となる。一方、完全 ADC で疑問値をそのまま真の異常値として処理する場合は、もともと正しかったデータを異常値として棄ててしまう危険をさけるため、適中率にウェイトを置き、疑問値検出率を比較的小きくすることとなる。

前者の場合の疑問値検出率の実際的な数値は、ADC の重要な目的の一つが人手によるチェック作業を大幅に軽減するということにあるので、人手によるチェック必要データ量がもとのデータ数より大幅に減少するのでなければ意味がない。このことからすると、部分 ADC における疑問値検出率は 0.1 以下（大きくても 0.2 以下）であるべきであると言える。また、別の視点から、統計学的な 1 つの目安として、 $3\sigma$  または  $2\sigma$  の範囲外の値を疑問値としての対象にすることが考えられる。すなわち、

$$(\bar{e}_j - 3\sigma) \leq e_j \leq (\bar{e}_j + 3\sigma)$$

ここで、 $e_j$  は予測値と実測値との差、 $\bar{e}_j$  は  $e_j$  の平均値、 $\sigma$  は  $e_j$  の標準偏差、上式の範囲に  $e_j$  が入らないとき当該データを疑問値とする。このときの疑問値検出率は、 $e_j$  の分布が正規分布であるとした場合、全体のデータ数に対して、約 0.3 % であり、上記の式で  $3\sigma$  のかわりに  $2\sigma$  を用いると、疑問値検出率は約 5 % となる。したがって 0.3 ~ 5 % 程度の疑問値検出率が一つの目安である。図 5 で見ると、疑問値検出率が 2 ~ 5 % 程度だと発見率も適中率も比較的高く維持できるが、疑問値検出率がたとえば 1 % 以下では発見率が 0.2 以下と低い値になり、10 % 以上では適中率が 0.2 以下と低い値になる。したがって、疑問値検出率は少なくとも 1 ~ 10 % の範囲外では実用的とはいえない。

リアルタイム・チェックで完全 ADC を考えるならば、すでに述べたように、発見率よりも適中率にウェイトをおく方がよく、疑問値検出率を小さ目にとる



ことがよいので、上記の範囲のうち小さい側（１～２％程度）が適当とみられる<sup>2)</sup>。一方、検出された疑問値を手でチェックし判断する部分 ADC の場合は発見率にウェイトをおくこととなるので、疑問値検出率を大きく、すなわち上記の範囲のうち大きい側（４～１０％）程度が適当とみられる<sup>2)</sup>。

### 3.4 ADF 法と加重平均法、偏差法との比較

ADF によるデータチェック法（以下、ADF 法と呼ぶ。）と加重平均法、偏差法を比較検討する。ADF 法は過去の観測値のみを対象とする方法であり、加重平均法は過去の観測値と未来の観測値を対象とする方法である。未来の観測値（既知であるとする。）も用いることができればデータの予測の精度は通常上がる。しかし、リアルタイム・チェックには用いられない。偏差法は最も簡易な方法であり、かなり広く用いられている。ADF 法が偏差法、加重平均法と比べて、その有効性が同程度であれば、この方法はオフライン・チェックのみならずリアルタイム・チェックにも利用するという長所があるので実用上の有効性が証明される。

#### ① 比較対象の手法

##### 加重平均法

測定値を時刻順に  $x_1, x_2, \dots, x_N$  とするとき、 $x_j$  の推定値  $\hat{x}_j$  を次式で求める。

$$\hat{x}_j = \sum_{\ell=1}^{2p+1} w_{\ell} x_{j-p+\ell-1}$$

ただし、 $2p+1 < N, p < j \leq N-p$ ,

$w_{\ell}$  は加重平均のウェイトで、

$\ell$  と  $p$  によって特定の値が定ま

っている。これを表 2 に示した。

ここでは加重平均の次数  $p$  は 3,

5, 7 を採用した。

表 2 加重平均のウェイト

$\ell$	$w_{\ell}$		
	P = 3	P = 5	P = 7
$\pm 7$			- 0.0832
$\pm 6$			- 0.0139
$\pm 5$		- 0.1059	0.0448
$\pm 4$		0.0265	0.0928
$\pm 3$	- 0.1429	0.1294	0.1301
$\pm 2$	0.2143	0.2029	0.1567
$\pm 1$	0.4286	0.2471	0.1727
0	0.0	0.0	0.0

## 偏差法

測定値を時刻順に  $x_1, x_2, \dots, x_N$  とするとき、 $x_j$  の推定値  $\hat{x}_j$  を次式で求める。

$$\hat{x}_j = x_{j-1}$$

## ② 比較実験

用いたデータは前項で用いたデータと同じである。SO<sub>2</sub> とオキシダントのデータを対象とした。まず、疑問値検出率と発見率の関係を SO<sub>2</sub> のデータを用い、ADF 法、加重平均法、偏差法のそれぞれについて調べた。<sup>8)</sup> ADF 法は Non-recursive  $p=10$  , Non-recursive  $p=24$  , Recursive  $p=24$  の3種類、加重平均法は次数  $p=3, 5, 7$  の3種類をとりあげた。

図 6 (1) は偏差法と3種類の加重平均による手法とを比較したもの、図 6 (2) は、偏差法と3種類の ADF による手法を比較したものであり、横軸の検出個数は対数で表わしてある。この例の場合、実際の異常値の個数 ( $N_0$ ) は 63 である。図において、上方に位置するものほど有効な手法

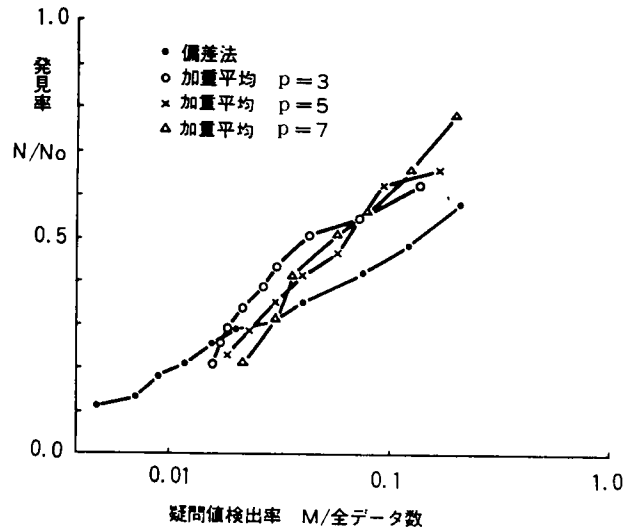


図 6 (1) データチェック手法による発見率(1)

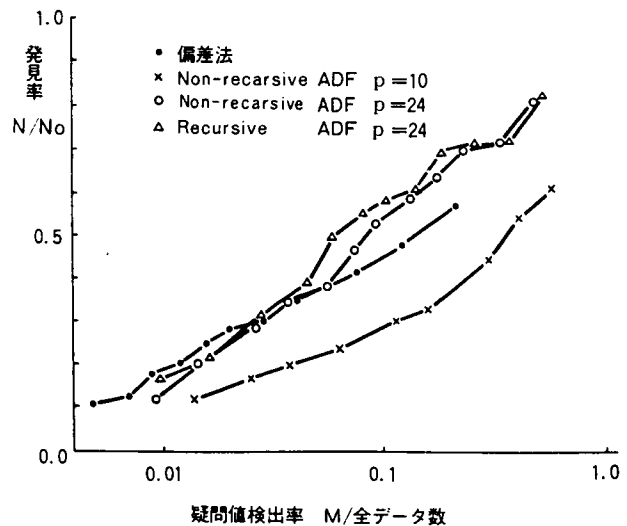


図 6 (2) データチェック手法による発見率(2)

であるので、加重平均値法（ $p = 3, 5, 7$ ）及び ADF 法（ $p = 24$ ）は、偏差法と比較して有効な手法であると言える。また図において見られるように、偏差法は、カーブの傾きがゆるやかであるので、疑問値の検出数（ $M$ ）を増やしてもそれほど発見率（ $N/N_0$ ）は上昇しないが、加重平均法や ADF 法（ $p = 24$ ）では傾きが比較的急であるので、 $M$ を増加させることによりかなり高い発見率（ $N/N_0$ ）の上昇を得ることができる。図から、加重平均法は次数によって有効性はそれほど大きく変化しないこと、および ADF 法では、Recursive モデル  $p = 24$ 、Non-recursive モデル  $p = 24$  のものが最も有効であることが読みとれる。

他のケースについてもテストしたが、その結果を表 3 に一覧表とした。<sup>8)</sup>

表 3 異常値検出手法の評価（各手法による適中率一覧表）

	異常値 の数 ( $N_0$ )	発見率 ( $N/N_0$ )	偏差法	加 重 平 均 法			ADF 法*				
				P - 3	P - 5	P - 7	Non 10 $u, v = 5 \times 10^{-7}$	Non 24 $u, v = 5 \times 10^{-7}$	Rec 24 $u, v = 5 \times 10^{-7}$	Rec 24 $u, v = 10^{-6}$	Rec 24 $u, v = 2 \times 10^{-6}$
SO <sub>2</sub> ケース 1	29	30% 40 50	2.2% 2.1 2.7	8.1% 9.2 3.8	3.8% 2.6 2.1	5.5% 2.6 1.9	2.2% 2.0 1.8	2.6% 2.3 2.5	2.9% 2.9 2.4		
ケース 2	63	30 40 50	40.2 16.5 10.1	44.0 39.4 35.8	34.4 30.4 17.5	32.0 32.7 28.4	7.7 4.4 4.7	25.9 20.0 19.9	42.0 33.6 23.5		
ケース 3	24	30 40 50	5.6 4.3 4.2	5.7 5.5 5.7	7.5 7.2 7.1	6.8 6.6 4.7	2.4 2.6 2.3	3.0 2.3 2.4	3.2 3.6 3.0		
Ox ケース 4	38	30 40 50	19.3 14.3 7.8	28.5 18.3 12.7	27.8 19.0 7.2	31.7 14.5 13.5	6.7 4.3 4.1	8.8 5.3 3.3	8.0 5.9 3.3	12.5% 8.4 5.4	
ケース 5	81	30 30 50	29.3 19.5 10.5	42.6 27.9 14.9	25.6 18.3 14.7	15.4 14.5 9.8	15.7 9.9 9.6	14.3 10.5 7.6	12.9 7.6 6.6	16.2 9.8 7.2	18.8% 10.7 7.9

\* Non 10, Rec 24 等の Non は non-recursive モデルを表わし 10, 24 はそれぞれ ADF の次数を示す。

この表ではそれぞれのケースについて、発見率（ $N/N_0$ ）が 30%、40%、50%のときの各手法による適中率（ $N/M$ ）を示した。表からわかるように、ケースによって異常値の検出結果にかなりの差があるが、全体として見ると

ADF法の有効性は加重平均法、偏差法とほぼ同程度かケースにより若干低いと判断できる。

### 3.5 項目間回帰法、上下限值法の有効性の検討

時系列値チェック法に別のチェック法を併用すると有効性が向上することが考えられる。併用すべきチェック法として項目間回帰法、上下限值法を取りあげる。なお、ここでは時系列チェック法として加重平均法を採用する。

#### (1) 検討対象手法

##### 項目間回帰法

異常値を検出しようとする測定項目の測定値を時刻順に

$$x_1, x_2, \dots, x_N$$

とし、他の $\ell$ 個の測定項目の測定値を時刻順に、

$$y_{11}, y_{12}, \dots, y_{1N}$$

$$y_{21}, y_{22}, \dots, y_{2N}$$

$$\vdots$$

$$y_{\ell 1}, y_{\ell 2}, \dots, y_{\ell N}$$

とする。 $y_{kj}$ は項目 $k$ 、時刻 $j$ についての測定値である。 $x_j$ の推定値 $\hat{x}_j$ を $y_{1j}, y_{2j}, \dots, y_{\ell j}$ から重回帰法により求める。時間帯による影響を加味する場合としない場合があるが、ここでは、時間帯による影響を加味して次の変数を設けた。すなわち、1～6時、7～12時、13～18時、19～24時の4つの時間帯について対応の変数を $T_1, T_2, T_3, T_4$ とし、時刻 $j$ が

$$1 \sim 6 \text{ 時のときは } T_{1j} = 1, T_{2j} = T_{3j} = T_{4j} = 0,$$

$$7 \sim 12 \text{ 時のときは } T_{2j} = 1, T_{1j} = T_{3j} = T_{4j} = 0,$$

$$13 \sim 18 \text{ 時のときは } T_{3j} = 1, T_{1j} = T_{2j} = T_{4j} = 0,$$

$$19 \sim 24 \text{ 時のときは } T_{4j} = 1, T_{1j} = T_{2j} = T_{3j} = 0,$$

とする。 $\hat{x}_j$ は次式で求められる。

$$\hat{x}_j = \alpha_1 y_{1j} + \alpha_2 y_{2j} + \dots + \alpha_\ell y_{\ell j} + \beta_1 T_{1j} + \beta_2 T_{2j} + \beta_3 T_{3j} + \beta_4 T_{4j}$$

係数 $\alpha_1, \alpha_2, \dots, \alpha_\ell, \beta_1, \beta_2, \beta_3, \beta_4$ は、

$$\sum_{j=1}^N (\hat{x}_j - x_j)^2$$

が最小となるように定める。

上記の式より推定値  $\hat{x}_j$  が求まれば、これと実測値との差  $e_j$  を計算し、

$C_L \leq e_j \leq C_U$  の範囲外のことを疑問値として検出する。ここでは、 $C_L$ ,  $C_U$  の値に  $3\sigma$  相当値を与えた。すなわち、 $C_L = \bar{e} - 3\sigma$ ,  $C_U = \bar{e} + 3\sigma$  である。 $\sigma$  は  $e_j$  の標準偏差で次の式から求められる。

$$\sigma = \sqrt{\frac{1}{N} \sum_{j=1}^N (e_j - \bar{e})^2}$$

$$\bar{e} = \frac{1}{N} \sum_{j=1}^N e_j$$

#### 上下限值法

上限値  $C_U$  , 下限値  $C_L$  を与えて、 $C_L$  と  $C_U$  の間に入らないデータがあると、そのデータを疑問値とする方法である。 $C_U$  ,  $C_L$  の値は原理的に、または實際上起こり得ない値やデータの度数分布からみて起こりにくい値（たとえば  $3\sigma$ （ $\sigma$  は標準偏差）相当値）を採用する。ここで用いた方法では、 $C_U = 8.0$ （ $0.8 \text{ ppm}$  または  $0.8 \text{ mg/m}^3$  の意味）、 $C_L = 1$ （ $0.01 \text{ ppm}$  または  $0.01 \text{ mg/m}^3$  の意味）を採用した。

#### 加重平均法

3.2(3)で述べたのと同じ方法である。次数  $p$  は 7 を採用した。この方法で求められた予測値と実測値との差  $e_j$  を計算し、 $C_L \leq e_j \leq C_U$  の範囲外のことを疑問値として検出する。 $C_L$  ,  $C_U$  の値は、ここでは  $3\sigma$  相当値を与えた。すなわち、 $C_L = \bar{e} - 3\sigma$ ,  $C_U = \bar{e} + 3\sigma$  である。

#### (2) 実データによる検討<sup>9),10)</sup>

用いたデータは大分県の昭和 52 年度の測定データのうち、4 月、7 月、10 月、12 月のデータであり、2 測定局（A 局、B 局）、5 測定項目（ $\text{SO}_2$  ,  $\text{NO}_x$ ,  $\text{NO}$ ,  $\text{O}_x$ , 浮遊粒子状物質）の分である。対象データはこのようなデータのケースの組合せの中から任意にいくつかのケースを選んだ。

加重平均法による疑問値検出の実験では、6 ケースについて調べたが、この結果は表 4 に示すように  $N/N_o$  は 0.11 ～ 1.0、 $N/M$  は 0.33 ～ 1.0 であった。

表 4 加重平均法、上下限值法による発見率と適合率

ケース	加重平均法		上下限值法		併用法(注)	
	$N/N_o$	$N/M$	$N/N_o$	$N/M$	$N/N_o$	$N/M$
SO <sub>2</sub>	0.29	0.33	0.57	1.00	0.71	0.18
NO <sub>x</sub> (その1)	0.11	0.78	0.34	0.95	0.44	0.90
(その2)	0.27	0.50	0.91	0.37	1.00	0.35
NO	0.24	1.00	0.37	0.40	0.50	0.48
O <sub>x</sub>	0.75	0.60	0.50	0.17		
PM	1.00	0.38	1.00	0.25		

(注) 上下限值法と加重平均法の併用法

上下限值法についての実験では、同様に 6 ケースについて調べたが、この結果は表 4 に示すように  $N/N_o$  は 0.34 ～ 1.0、 $N/M$  は 0.17 ～ 1.0 で、加重平均法とはほぼ同程度の値であり、かなり有効性が高いと考えられる。

項目間回帰法についての実験では、9 ケースについて調べたが、この結果は表 5 に示すように非常に低い値であった。 $N/N_o = 0$ 、 $N/M = 0$  の場合が 4 ケースもあった。また、重相関係数は表 6、表 7 に示すように低かった(高い場合でも 0.7 ～ 0.8、低い場合は 0.2 ～ 0.3)。したがって、この方法は有効性は高くないと評価される。

表 5 項目間回帰法による発見率と適合率

ケース	$N/N_o$	$N/M$
SO <sub>2</sub> (その1)	0.14	0.09
(その2)	0	0
(その3)	0.02	0.13
(その4)	0	0
(その5)	0	0
(その6)	0.5	0.07
NO <sub>x</sub> (その3)	0	0
(その4)	0.01	0.33
NO	0.12	1.0

次に、加重平均法と上下限值法を併用するどのように有効性が向上するかをみる。表 4 は、加重平均法、上下限值法のそれぞれによる  $N/N_o$ 、 $N/M$  とその併用法によるそれとを表示したものである。併用法により発見率( $N/N_o$ )がかなり高められることがわかる。

表6 各項目間の相関係数と重相関係数（A測定局のデータ）

データの種別			SO <sub>2</sub>	NO <sub>x</sub>	NO	O <sub>x</sub>	P.M.
4月のデータ	相 関 係 数	NO <sub>x</sub>	0.29				
		NO	- 0.01	0.18			
		O <sub>x</sub>	0.35	- 0.08	- 0.04		
		P.M.	0.26	0.46	0.09	- 0.28	
	重相関係数(注)		0.54	0.51	0.19	0.52	0.57
7月のデータ	相 関 係 数	NO <sub>x</sub>	0.24				
		NO	0.17	0.66			
		O <sub>x</sub>	0.53	0.08	- 0.10		
		P.M.	0.02	0.12	- 0.01	0.16	
	重相関係数(注)		0.59	0.69	0.69	0.59	0.24
10月のデータ	相 関 係 数	NO <sub>x</sub>	0.14				
		NO	0.07	0.61			
		O <sub>x</sub>	0.39	- 0.15	- 0.09		
		P.M.	0.18	0.02	- 0.08	0.25	
	重相関係数(注)		0.45	0.64	0.61	0.48	0.29
12月のデータ	相 関 係 数	NO <sub>x</sub>	0.37				
		NO	0.26	0.67			
		O <sub>x</sub>	0.06	- 0.33	- 0.28		
		P.M.	0.36	0.40	0.38	- 0.36	
	重相関係数(注)		0.51	0.72	0.68	0.49	0.56

(注) 重相関係数は該当項目を他の項目から同補した場合の係数である。

表7 各項目間の相関係数と重相関係数（B測定局のデータ）

データの種別			SO <sub>2</sub>	NO <sub>x</sub>	NO	O <sub>x</sub>	P.M.
4月のデータ	相 関 係 数	NO <sub>x</sub>	0.30				
		NO	0.05	0.56			
		O <sub>x</sub>	- 0.03	- 0.26	- 0.23		
		P.M.	0.25	0.58	0.53	- 0.30	
	重相関係数(注)		0.37	0.69	0.63	0.33	0.66
7月のデータ	相 関 係 数	NO <sub>x</sub>	0.33				
		NO	0.25	0.31			
		O <sub>x</sub>	0.14	0.30	- 0.16		
		P.M.	0.32	0.44	- 0.01	0.27	
	重相関係数(注)		0.43	0.60	0.46	0.42	0.52
10月のデータ	相 関 係 数	NO <sub>x</sub>	0.50				
		NO	0.18	0.63			
		O <sub>x</sub>	0.41	0.11	- 0.16		
		P.M.	0.35	0.36	0.37	- 0.07	
	重相関係数(注)		0.66	0.74	0.69	0.50	0.50
12月のデータ	相 関 係 数	NO <sub>x</sub>	0.61				
		NO	0.47	0.50			
		O <sub>x</sub>	0.04	0.09	0.08		
		P.M.	0.44	0.39	0.71	- 0.05	
	重相関係数(注)		0.65	0.65	0.76	0.18	0.73

#### 第4節 結 語

環境汚染の測定データには各種の原因で異常値や誤差が混入することは避けられず、これをどのようにチェックするかはデータの精度の向上の観点から重要である。本章では、大気環境データをとりあげ、現在全国の自治体で採用されてい

るデータチェック方法が有する問題点とこれを解決するための方法——コンピュータを用いたデータの自動チェックシステムの採用——を指摘した。そして、データ自動チェックシステムにおいてどのようなチェック手法を採用するのが有効かについて検討した。

データチェック手法の有効性を評価するための判断尺度として、発見率( $N/N_0$ )と適中率( $N/M$ )の2つの量を定義し、これらの値により各手法の有効性を比較検討した。得られた結果の概要は次のようである。

- ① ADF ( Adaptive Digital Filter ) 法をデータチェック法に用いることを検討した。まず、ADF によればデータ値は過去のデータを用いてかなり正しく予測されることを示した。予測値と実測値との差の大きいデータを疑問値として検出し、それがどの程度真の異常値であるか、また検出しえなかった異常値がどの程度あるかを、疑問値検出率に応じた発見率及び適中率を算出することにより検討した。
- ② ADF 法によるデータチェック法では疑問値検出率が0.02～0.04のとき発見率、適中率が0.3～0.4となる。疑問値検出率が増大するにつれ発見率は上がるが、一方適中率は下がる。疑問値検出率はデータチェックの目的に応じて適当な値を設定すればよい。実用上の観点から0.01～0.1程度が適当である。
- ③ 時系列値チェック法 ( ADF法、加重平均法、偏差法 ) についてその有効性を相互に比較した。ADF法の有効性は加重平均法、偏差法とほぼ同程度であるかケースにより若干低い。
- ④ データの予測を過去の観測値群のみで ( 未来の観測値は用いないで ) 行う ADF 法によるデータチェック法は、リアルタイム・チェックに用いることが可能である。
- ⑤ 複数のデータチェック手法を併用することによる有効性の向上について検討した。時系列値チェック法と併用する手法として上下限值法は有効性が高いが、項目間回帰法は低い。加重平均値法に上下限值法を併用することによりかなり発見率を高めることができる。



## 第6章 データの2次元化手法に関する研究<sup>1) 2)</sup>

### 第1節 概 説

大気汚染や水質汚濁などの測定データは、平面上の離散的な数の限られた地点について得られるのが通常である。一方、環境汚染の評価及び対策においては、次の①～④の目的にデータを利用するので、汚染状況を2次元の分布パターンとして把握することが重要である。すなわち、

- ① 地域または水域全体の大気汚染状況や水質汚濁状況を推定する。
- ② 汚染分布の状況を視覚的に把握する。
- ③ 2次元メッシュ点上の汚染濃度値を推定し、メッシュデータで表わされた他の関係データとの間での関連解析等を行う。

離散的な濃度データから連続的な濃度分布を推定するためのデータ補間法はいく種類か知られているが、<sup>3)</sup>そのひとつとしてスプライン法がある。

渡津<sup>4)</sup>は、スプライン法によって作成された大気汚染の補間曲面と大気拡散式を用いて予測した濃度分布曲面とを比較することにより、スプライン法によるデータ補間法の精度を論じている。また、篠原<sup>5)</sup>は、大気汚染測定局の実測データの一部だけを用いてスプライン補間を行い、その計算結果と残りのデータとを比較することにより、この誤差を最小にするスプライン・パラメータの値を論じている。

本章では、大気汚染近似分布の、大西の<sup>6)</sup>スプライン法にもとづく形成法と、必要な程度の近似になっているか否かの判定方法を提案する。また、この手続により、近似分布を得るに必要な測定局が同時に求められる。さらに、大気汚染の実データに提案した形成法を適用して、妥当な近似曲面がえられることを示す。

### 第2節 スプライン法

#### 2.1 特 徴

従来から用いられている分布パターンを得るための各種手法のうち、多項式補間については、データ点の疎密に応じて各格子点の代表領域を変えねばならない

難点があり、また、高次の多項式やフーリエ級数などで補間する方法は直観に合わない結果がしばしば見られるなどの問題がある<sup>6)</sup>。また、多項式を用いる代りに、距離に応じた重みつき平均をとる方法もある<sup>6)</sup>。どの方法もデータがかなり密に存在することが必要であり、また、外挿して推定された境界付近の濃度は極端に不自然なものになることがあるなど、種々の問題点を有している。

これに対して、スプライン法は、測定局密度が小さくても、スプライン・パラメータによって滑らかな曲面を得ることが可能であり、また、測定局での測定値がそのまま補間分布曲面の値に採用されるなど環境汚染データの補間法として望ましい特徴を有している<sup>5)</sup>。

スプライン法の直観的原理は以下のようなものである。平面上のいくつかの点において与えられたある値（たとえば汚染物質の濃度）に対応する高さの棒が何本か立っているとする。その上にある弾性をもつ一枚の板をかぶせ、板が棒の先端にふれるようにして、その板が形成する曲面形状をもって、全域の濃度分布形を表わそうとするものである<sup>5)</sup>。

## 2.2 計算方法

この方法の定式化と差分方程式へのおきかえについて、大西<sup>6)</sup>にもとづいて述べる。

座標軸（ $x$ ， $y$ ， $z$ ）を次のようにとる。2次元水平面（ $x$ ， $y$ ）上のデータ点におけるデータの値に相当する高さを $z$ とする。データ点（ $x$ ， $y$ ）＝（ $x_i$ ， $y_i$ ），（ $i = 1, 2, \dots, N$ ）において、それぞれ $z = z_i$ ，（ $i = 1, 2, \dots, N$ ）となるような曲面 $z = Z(x, y)$ を求めることを考える。弾性体の微小変位理論において、歪エネルギーと張力による伸びのエネルギーの和に比例する量として、(1)式が定義される。

$$E = \iint_A \{ (\Delta Z)^2 + \sigma (\nabla Z)^2 \} dx dy \dots\dots\dots (1)$$

$$\left( \text{ただし、} \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad \nabla = \frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} \right)$$

$\sigma$  は弾性板の張力による伸びのエネルギーと歪エネルギーの係数の比を表わして

いる。(1)式の積分の範囲  $A'$  は、対象とする領域  $A$  を含む適当に広い領域とする。 $A'$  の境界  $\partial A'$  が、対象とする領域  $A$  より、十分遠くにとられているならば、境界条件は、

$$\frac{\partial Z}{\partial n} = \Delta Z = 0 \quad \text{on } \partial A' \dots\dots\dots(2)$$

(ただし、 $\partial/\partial n$  は  $\partial A'$  の外向き法線の方向への微分) とすればよい。

(1)式の変分を 0 とおくと

$$\Delta^2 Z - \sigma \Delta Z = 0 \dots\dots\dots(3)$$

すなわち、 $Z(x, y)$  はデータ点においては、データ値をとり、 $A$  において

(3)式を満たし、境界  $\partial A'$  において、条件(2)式を満たすような関数として定まる。

実際の計算のために、(3)式を差分方程式におきかえる。

簡単のため、領域  $A'$  内に、正方形格子

$$\left. \begin{aligned} x &= X_\alpha \quad (\alpha=1, 2, \dots, p), \\ y &= Y_\beta \quad (\beta=1, 2, \dots, q), \\ X_\alpha - X_{\alpha-1} &= Y_\beta - Y_{\beta-1} = h, \end{aligned} \right\} \dots\dots\dots(4)$$

を考える。各格子点  $(X_\alpha, Y_\beta)$  に対して、代表領域  $r_{\alpha, \beta}$  を

$$\left. \begin{aligned} X_\alpha - \frac{h}{2} &\leq x < X_\alpha + \frac{h}{2}, \\ Y_\beta - \frac{h}{2} &\leq y < Y_\beta + \frac{h}{2}, \end{aligned} \right\} \dots\dots\dots(5)$$

と定義する。

(3)式は、次の 2 つの差分方程式におきかえられる。メッシュ  $r_{\alpha, \beta}$  の  $Z$  を  $Z_{\alpha, \beta}$  とする。

$$\left. \begin{aligned} (Z_{\alpha-1, \beta} + Z_{\alpha+1, \beta} + Z_{\alpha, \beta-1} + Z_{\alpha, \beta+1} - 4Z_{\alpha, \beta})/h^2 &= V_{\alpha, \beta} \\ (V_{\alpha-1, \beta} + V_{\alpha+1, \beta} + V_{\alpha, \beta-1} + V_{\alpha, \beta+1} - 4V_{\alpha, \beta})/h^2 &= \sigma V_{\alpha, \beta} \end{aligned} \right\} \dots\dots\dots(6)$$

ただし、 $r_{\alpha, \beta}$  がデータ点  $(x_i, y_i)$  を含む場合には、補間曲面がデータ点を通ることを表わす式 (例えば(7)式)

$$z_i = Z_{\alpha, \beta} + \frac{Z_{\alpha+1, \beta} - Z_{\alpha, \beta}}{h} (x_i - X_{\alpha}) + \frac{Z_{\alpha, \beta+1} - Z_{\alpha, \beta}}{h} (y_i - Y_{\beta}) \dots\dots\dots(7)$$

で(6)式をおきかえる。または、データ点が  $r_{\alpha, \beta}$  の中心点にあると見なして、そのデータ点を通ることを表わす式でおきかえる。結局、(6)式または(7)式から逐次近似法により解を求め、格子点での値を得る。

## 2.3 計算上の条件

(1)式の計算の範囲  $A'$  を長方形にとった。格子、境界条件、初期値は次のとおりとした。

代表領域の幅を  $h = 1 \text{ km}$  とした。代表領域（以下、メッシュと呼ぶ。）に1つのデータが存在するとき、それを、メッシュに与えるデータとして採用した。複数のデータが存在するときは、その算術平均を採用した。

境界条件として、境界における1次、2次の微分係数を0とした。つまり、 $Z_{1, \beta} = Z_{3, \beta}, Z_{2, \beta} = Z_{3, \beta}$ （ただし、ここで  $\beta = 3, 4, \dots, q-2$ ）である。 $Z_{\alpha, 1}, Z_{\alpha, 2}$  についても同様である。

領域  $A'$  の四囲から、境界条件を与えるためのメッシュの2列分を除いた長方形領域を  $A$  とする。

スプライン法による計算を開始するとき、全格子点（境界各2列も含む）に次の方法により初期値  $Z_{\alpha, \beta}^0$  を与えた。データ点が存在するメッシュにおいては、

$$Z_{\alpha, \beta}^0 = \sum_{i=1}^a z_i / a$$

（ただし、 $z_i$  はメッシュ  $r_{\alpha, \beta}$  に含まれるデータの値、 $a$  はその個数）

データ点が存在しないメッシュにおいては、

$$Z_{\alpha, \beta}^0 = \{ (z_i)_1 \times d_2 + (z_i)_2 \times d_1 \} / (d_1 + d_2)$$

（ただし、 $(z_i)_1, (z_i)_2$  は格子点  $(X_{\alpha}, Y_{\beta})$  に最も近い2つのデータ点のデータの値、 $d_1, d_2$  は格子点からそれぞれのデータ点までの距離の2乗、なお、最も近いデータ点が3個以上存在するときは、データ点につけた番号の若いものから2個をとるとした。）

逐次近似の計算では、次の式が成立するとき収束したものとみなした。

$$\text{Max}_{(X_{\alpha}, Y_{\beta}) \in A} |Z_{\alpha, \beta}^i - Z_{\alpha, \beta}^{i+1}| < 10^{-3}$$

(ただし、 $Z_{\alpha, \beta}^i$  は、繰返し回数  $i$  回目のときの格子点  $(X_{\alpha}, Y_{\beta})$  の値を示す。)

なお、この収束判定条件を満足しなくても、繰返し回数  $i$  が 500 となったら、計算を打ち切った。

## 2.4 パラメータ $\sigma$ の決定方法

スプライン法を用いる場合、パラメータ  $\sigma$  の値によって、補間分布曲面の形状が異なる。たとえば、 $\sigma = 0$  の場合は、周辺からの張力がかけられない場合に弾性板が自然にとる形状を意味する。この場合、非常に滑らかな曲面形状となるが、直観に合わないような極値が現われたりする。一方、 $\sigma = \infty$  の場合は、データ点を通る膜の平衡状態の形状を表わす。この場合、データ点で尖りのある不自然な曲面形状となる。

適当なパラメータの決定については、 $\text{SO}_2$  の実測データを補間により求めたデータと比較することにより調べた篠原ら<sup>5)</sup>の研究がある。これによると、 $\sigma = 1$  近辺がよい結果を与えるとしているが、 $\sigma = 0$  では  $\sigma = 1$  よりかなり結果が悪く、 $\sigma = 1 \sim 10$  程度では概ね  $\sigma = 1$  と同程度の結果を与えるとしている。

本文では、次のような方法で  $\sigma$  の最適値を調べる。

- ① 領域  $A$  に存在する  $N$  個の測定局から、ランダムに 1 局を取除いて、補間曲面を作る。取除かれた局  $i$  の測定値を  $z_i$ 、補間によって得られた推定値を  $\hat{z}_i$  とする。 $N$  局からランダムに 1 局取除く操作を  $J$  回繰返し、次の値を求める。

$$y = \frac{\sum_{i \in \Omega} (\hat{z}_i - z_i)^2}{\sum_{i \in \Omega} (z_i - \bar{z})^2}$$

(ただし、 $\Omega$  は取除いた局の集合で、 $\Omega$  の要素の数は  $J$ 、また、 $\bar{z} = \frac{\sum_{i \in \Omega} z_i}{J}$ )

- ② 異なるいくつかの  $\sigma$  の値について、同じ  $\Omega$  を用いて①の計算を行い、 $y$  が小となる  $\sigma$  を選ぶ。

### 第3節 補間の適用性の検証方法

スプライン法で汚染濃度が必要な程度に補間できているか否かの議論は、厳密には真の濃度分布がわからないと行えない。しかし、真の濃度分布は知りえないので、真の濃度分布に十分近いと推定される近似濃度分布（以下、これを近似分布という。）を得ることを考える。

本節では、近似分布のスプライン法による形成法と、必要な程度の近似になっているか否かの判定法を提案する。

#### 3.1 近似分布の決定

いま、 $x-y$  平面上の領域  $A$  における真の濃度分布曲面を  $P(x, y)$  とし、対象領域  $A$  には  $N$  個の測定局が存在するとする。 $s$  を  $N$  局から  $N$  局を選び出す順列のひとつとする。 $i$  番目の順列と特定する場合には、 $s(i)$  と記す。 $n$  を全ての順列のうち、ランダムに選んだ順列の集合、また、その要素数とする。 $P_k^s(x, y)$  を  $s$  なる順列のうち  $k$  番目までの測定局データをスプライン法で補間した曲面とする。さらに、 $P_k^s(x, y)$  の比較曲面を

$$P_k^{s*}(x, y) = P_{k+r}^s(x, y) \dots\dots\dots (8)$$

で定義する。ただし、 $r$  は 1 以上の整数、また、一般に関数  $F(x, y)$  の  $A$  上の平均値を  $[F]$  で表わす。

$$[F] = \frac{1}{A} \iint_A F(x, y) dx dy$$

スプライン法による補間分布曲面  $P_k^s(x, y)$  と、これとの比較曲面  $P_k^{s*}(x, y)$  のへだたりを、 $k$  を増大させつつ、評価する。へだたりの尺度を

$$D_k^s = \frac{1}{A} \iint_A |P_k^s(x, y) - P_k^{s*}(x, y)| dx dy \dots\dots\dots (9)$$

とする。

$D_k^s$  は、 $k \rightarrow$  大のとき、 $P_k^s(x, y)$  が  $P(x, y)$  に近づくなれば  $D_k^s$  は零に近づき傾向を持つ。なぜならば、もし、 $N$  が無限大であって、 $k$  が大になるとき、

$$\frac{1}{A} \iint_A |P_k^s(x, y) - P(x, y)| dx dy \rightarrow 0 \quad \cdots \cdots \cdots (10)$$

とする。一方、次の式がなりたつ。

$$D_k^s = \frac{1}{A} \iint_A |P_k^s(x, y) - P(x, y) - P_k^{s*}(x, y) + P(x, y)| dx dy \leq B_k^s + B_{k+r}^s \quad \cdots \cdots \cdots (11)$$

ただし、ここで

$$B_k^s = \frac{1}{A} \iint_A |P_k^s(x, y) - P(x, y)| dx dy \quad \cdots \cdots \cdots (12)$$

$$B_{k+r}^s = \frac{1}{A} \iint_A |P_{k+r}^s(x, y) - P(x, y)| dx dy \quad \cdots \cdots \cdots (13)$$

よって、(10)式が成立つとすれば、(11)式の右辺の2項のそれぞれが零に近づくので(9)式も零に近づく。

したがって、測定局が真の汚染濃度曲面  $P(x, y)$  の形状の特徴をとらえうるだけ密に存在していれば、 $N!$  個ある順列  $s$  の多くのものについては、 $D_k^s$  は  $k$  の増大につれて減少する傾向を持ち、零に近い値をとるであろう。

次に、 $P_k^{s*}(x, y)$  のそれ自身の平均値  $[P_k^{s*}]$  からのへだたりの程度を次式で定義する。

$$R_k^s = \frac{1}{A} \iint_A |P_k^{s*}(x, y) - [P_k^{s*}]| dx dy \quad \cdots \cdots \cdots (14)$$

$P_k^s(x, y)$  が、 $k$  の増大と共に、 $P(x, y)$  に近づくとすれば、 $R_k^s$  も一定値に近づく。なぜならば、 $N$  が無限大であって、 $k$  が大になるとき、

$$\frac{1}{A} \iint_A |P_k^{s*}(x, y) - P(x, y)| dx dy \rightarrow 0 \quad \cdots \cdots \cdots (15)$$

としている。一方

$$R = \frac{1}{A} \iint_A |P(x, y) - [P]| dx dy$$

とする。このとき

$$R_k^s - R$$

$$= \frac{1}{A} \iint_A |P_k^{s*}(x, y) - P(x, y) - [P_k^{s*}] + [P] + P(x, y) - [P]| dx dy - R$$

$$\leq \frac{1}{A} \iint_A |P_k^{s*}(x, y) - P(x, y)| dx dy + \frac{1}{A} \iint_A |[P_k^{s*}] - [P]| dx dy \dots\dots\dots (16)$$

なる関係が成立つ。(16)式の右辺第2項は

$$| \frac{1}{A} \iint_A \{ P_k^{s*}(x, y) - P(x, y) \} dx dy |$$

と変形できる。(15)式により、(16)式の右辺は、 $k$ が大になるにつれて、零に近づくので、 $R_k^s$ は一定値 $R$ に近づく。

近似分布は、次の考え方で定める。すなわち、ある順列 $s$ をとったとき、十分長い $k$ の区間 $[k', k'']$ に属する $k$ について、 $D_k^s/R$ が1より十分小なら、ある $k$ の値についてのみ偶然 $P_k^s(x, y)$ と $P_{k+r}^s(x, y)$ との曲面が近づいたため $D_k^s/R$ が小になるのではなく、(11)式における $B_k^s$ と $B_{k+r}^s$ が小になっているために、 $D_k^s/R$ が安定して小になっていると考える。 $R$ はもとより知りえないので、 $R$ に近づく値 $R_k^s$ をこれに代える。

結局、 $D_k^s/R_k^s$ が十分長い区間 $(k', k'')$ で小でありつづけると、測定局の追加が、分布曲面のそれ自身の平均値からのへだたり(分布曲面の $A$ 上での変化)にくらべ、もはや、 $P_k^s(x, y)$ の分布曲面をわずかししか改良しえないだけ、 $P_k^s(x, y)$ が $P(x, y)$ に近づいたと考える。

$n$ 個の順列のうち、 $i$ 番目の順列 $s(i)$ についての $D_k^s, R_k^s, P_k^s(x, y)$ をそれぞれ、 $D_k^{s(i)}, R_k^{s(i)}, P_k^{s(i)}(x, y)$ とする。

$n$ 個の順列 $s$ を考えたとき、 $D_k^{s(i)}$ は( $R_k^{s(i)}$ も)、 $s(i) \in n$ によって異なる。特定の $s(i)$ によって、測定局を増やしても、他の $s(i)$ に対して存在した十分長い区間 $(k', k'')$ が存在しないこともありうる。順列 $s(i)$ によっては、その終り近くなつてはじめて、分布曲面の特徴を支配している測定局があらわれることもあるからである。そのような順列を含みうる $n$ 個の順列を、 $N$ 個の測定局の順列の



すべて (N! 個ある) からランダムに選び、 $D_k^{s(i)} / R_k^{s(i)}$  の各 k での平均値を  $m_k$  , 標準偏差を  $s_k$  とする。

$m_k$  または  $m_k + s_k$  の値が小でありつづける十分長い区間  $[k', k'']$  を考える。このような区間が存在するとき、n 個あるうちの多くの  $s(i)$  について濃度分布  $P_k^{s(i)}(x, y)$ ,  $k \in [k', k'']$  は真の濃度分布  $P(x, y)$  に近いと推定され、なおかつ、測定局の設置数は  $k'$  個で不足がないものと推定される。

このとき、近似分布曲面  $P_a(x, y)$  は、次のものを採用すればよい。まず、 $s(i) \in n$  のうち、 $D_k^{s(i)} / R_k^{s(i)}$  の値が、 $m_k + s_k$  よりも、 $[k', k'']$  の全てで小さな値をもつ順列  $s(i)$  を選び出す。その  $i$  の集合を  $m$ 、集合の要素数も  $m$  とすると

$$P_a(x, y) = \frac{1}{mI} \sum_{i \in m} \sum_{k \in [k', k'']} P_k^{s(i)}(x, y) \dots\dots\dots (17)$$

(ただし、ここで、 $I$  は  $[k', k'']$  に属する整数の個数である。)

そのほか、近似分布曲面として次のようなものを採用してもよい。

$$P_a(x, y) = \frac{1}{m} \sum_{i \in m} P_k^{s(i)}(x, y) \dots\dots\dots (18)$$

$$P_a(x, y) = P_N(x, y) \dots\dots\dots (19)$$

(ただし、 $N \geq k''$  のときに限る。また、 $P_N(x, y)$  はすべての測定局データを採用してスプライン法で得た曲面である。

一方、このような近似分布が得られたならば、使用したスプライン補間法により、汚染濃度が十分補間され得たと推定される。

### 3.2 近似分布推定の妥当性

3.1 で述べた方法は、多変数関数の極値探索における探索打切——極値を与える点に到達したと見なすこと——に似ている。すなわち「いくら局数  $k$  (探索繰返し回数) を増やしても、もはや、 $D_k^s / R_k^s$  (目的関数値) が変化しないなら、真の濃度曲面 (極値を与える点) に近いものが得られているであろう。」と考えるのである。

このような考え方をする必要性は次のようである。すなわち、真の濃度分布  $P(x, y)$  (目的関数の真の最小値) が知りえないが故にこそ、それを推定する手続を構成しようとしているのであり、また真の分布 (真の最小値) が知りえないが故にこそ、真のものが得られたという絶対的検証ができないので、「真のものに近いものが得られたと推定する」という手続によらなければならないのである。

#### 第4節 実データへの適用

前節の方法の有効性を検討するために、この方法を実際の大気汚染測定データに適用する。有効性の検証には、起伏が顕著な濃度分布を用いるのが望ましいので、関東地域において、光化学スモッグが発生して、オキシダント濃度が高かった昭和49年8月3日のデータを採用した。この日の光化学スモッグの発生の継続時間 (午前10時頃から午後5時頃まで) の中から、汚染パターンの異なる時刻  $t_1 = 10$  時、 $t_2 = 12$  時、 $t_3 = 14$  時、 $t_4 = 16$  時を選び出した。濃度分布をスプライン補間する領域Aは、東西  $73.92\text{ km}$  (東経  $139^\circ 10' 0''$  から  $140^\circ 7' 30''$  まで)、南北  $46.25\text{ km}$  (北緯  $35^\circ 24' 0''$  から  $35^\circ 49' 30''$  まで) の長方形の領域で、この中に図1に示す  $N = 55$  の測定局がある。領域A' を、東西について  $45''$  ごとに  $66$  分割、南北について  $30''$  ごと

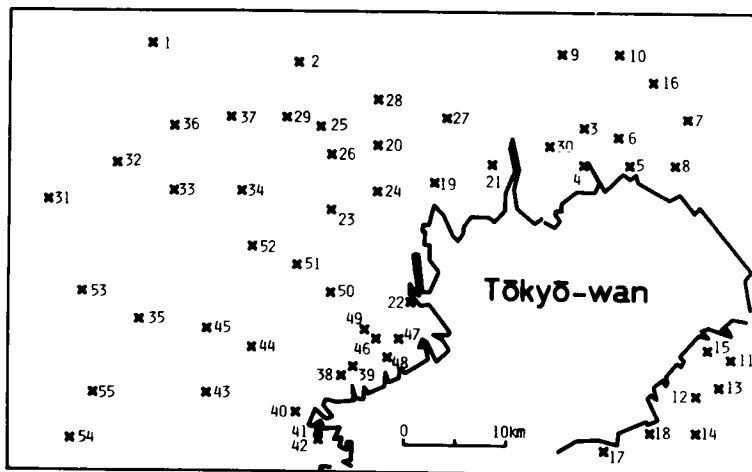


図1 スプライン法の適用性を検討する領域及び測定局

とに50分割してメッシュをつくった。この分割法は総務庁統計局の経緯度法による1キロメッシュの形成法にもとづいている。ひとつのメッシュの大きさは、厳密には、東西1.120 km、南北0.925 kmとなる。

#### 4.1 パラメータ $\sigma$ の決定

はじめに、2.4の方法に従って  $\sigma$  の値を決定する。

複数の分布曲面について  $\sigma$  の最適値を検討するため、上記データのうち、12時 ( $t_2$ ) と14時 ( $t_3$ ) におけるものを用いた。  $J = 17$  とした。図2は誤差  $y$  と  $\sigma$  の関係を表わす。ランダムに取除いた局の中にたまたま分布曲面の特徴を表わす重要な数局が含まれていたため、

$y$  の値が比較的大きくなった。  $\sigma = 1.0 \sim 10.0$  では  $y$  はほぼ一定である。  $\sigma$  の値として1.0から10.0の間のいずれかの値を採用すればよい。本章においては  $\sigma = 5.0$  を用いる。

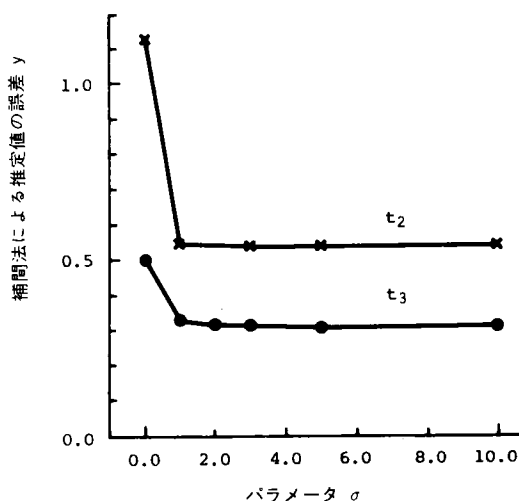


図2 パラメータ  $\sigma$  と補間法による推定値の誤差  $y$  との関係

#### 4.2 有効性の検討

(8)式の  $r$  として、  $r = 5$  ,  $r = 1$  を用いる。順列を  $n = 10$  個とったときの  $k$  における  $s(i)$  による  $D_k^{s(i)} / R_k^{s(i)}$  についての  $m_k$  と  $m_k \pm s_k$  を  $t_1 \sim t_4$  について、また  $r = 5$  ,  $r = 1$  について図3～図6に示す。

$m_k$  ,  $m_k + s_k$  の  $k$  の増大に伴う減少挙動について観察する。図3～図6において、ある値までの  $k$  の増大に対しては  $m_k$  ,  $m_k + s_k$  の減少挙動が顕著に見られ、その値以上の  $k$  においては、  $m_k$  ,  $m_k + s_k$  はほとんど一定の値で推移している。

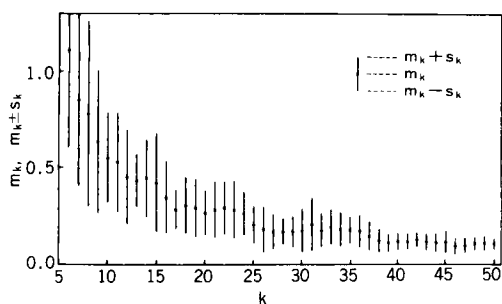


図 3 (a)  $k$  と  $m_k, m_k \pm s_k$  との関係  
(  $t = 10$  時,  $r = 5$  )

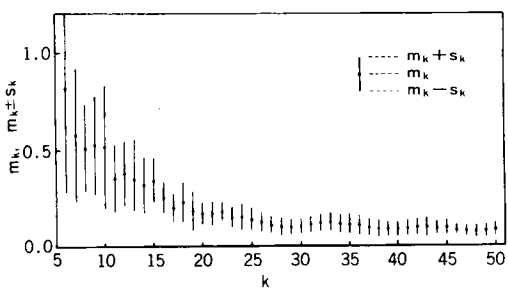


図 5 (a)  $k$  と  $m_k, m_k \pm s_k$  との関係  
(  $t = 14$  時,  $r = 5$  )

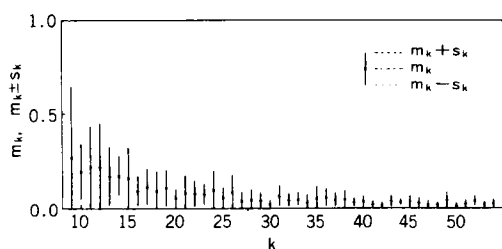


図 3 (b)  $k$  と  $m_k, m_k \pm s_k$  との関係  
(  $t = 10$  時,  $r = 1$  )

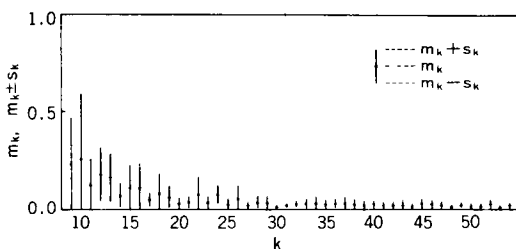


図 5 (b)  $k$  と  $m_k, m_k \pm s_k$  との関係  
(  $t = 14$  時,  $r = 1$  )

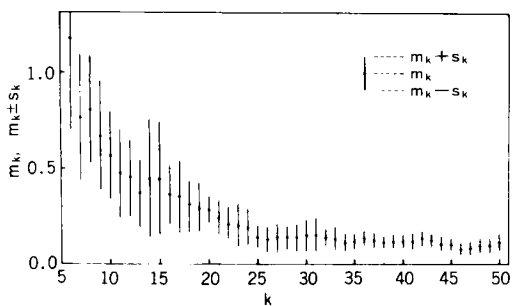


図 4 (a)  $k$  と  $m_k, m_k \pm s_k$  との関係  
(  $t = 12$  時,  $r = 5$  )

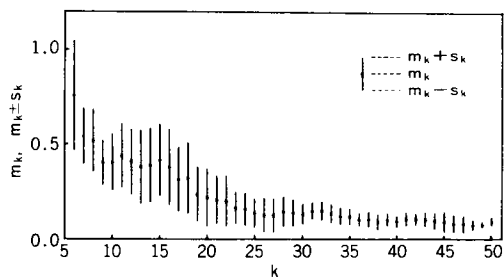


図 6 (a)  $k$  と  $m_k, m_k \pm s_k$  との関係  
(  $t = 16$  時,  $r = 5$  )

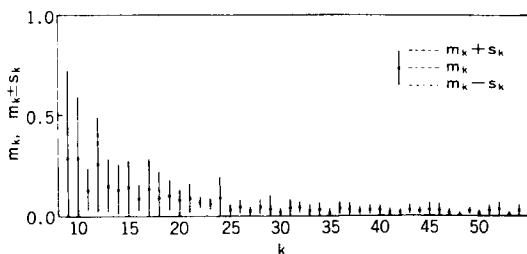


図 4 (b)  $k$  と  $m_k, m_k \pm s_k$  との関係  
(  $t = 12$  時,  $r = 1$  )

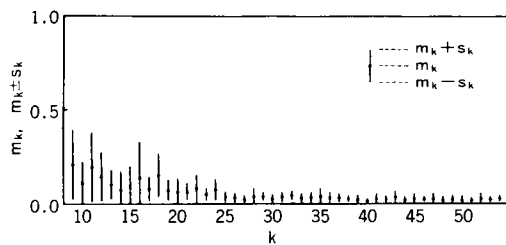


図 6 (b)  $k$  と  $m_k, m_k \pm s_k$  との関係  
(  $t = 16$  時,  $r = 1$  )

測定局のデータは必然的に測定誤差を持つ。 $m_k, m_k + s_k$  が、ほぼ一定値を持つ  $k$  の区間では、この誤差が支配的要因になっていると思われる。よって、ここでは、この区間において、 $m_k < c_m, m_k + s_k < c_s$  を満たす小さい概数  $c_m, c_s$  をとって、 $m_k, m_k + s_k$  が十分小かどうかを判断する値とする（表 1）。 $D_k^{s(i)}/R_k^{s(i)}$  の意味を考えれば、 $m_k < c_m, m_k + s_k < c_s$  を満たす  $k$  の値だけの個数の測定局があれば、濃度分布の A 上の変化の程度にくらべて、ある程度まで良い精度の近似濃度分布が得られていると考えられる。実用上は、 $c_m, c_s$  の値は、個々の事例の目的に応じて設定されなければならない。

表 1  $c_m, c_s$  の値

	$c_m$	$c_s$
$r = 5$	0.15	0.3
$r = 1$	0.05	0.1

表 2 に、 $m_k < c_m$  または、 $m_k + s_k < c_s$  を満たす  $k$  の最小値（時刻  $t_i$  における分布曲面に対するこの値を  $k'_i$  と表わす。）を  $r = 5, r = 1$  について示す。4 通り求まる  $k'_i$  のうち、安全側にとって、最大のものを採用し、これを同表の最右欄に示した。

表 2  $m_k < c_m, m_k + s_k < c_s$  となる  $k$  の区間の最小値  $k'_i$

$t_i$	$r = 5$		$r = 1$		$m_{\max}(k'_i)$
	$m_k$	$m_k + s_k$	$m_k$	$m_k + s_k$	
10 時	37	32	37	37	37
12 時	25	24	25	25	25
14 時	25	19	27	27	27
16 時	25	23	25	25	25

一方、 $m_k, m_k + s_k$  の  $k$  に対する挙動において、減少挙動の顕著な部分と、ほとんど一定値で推移している部分との間で、 $m_k, m_k + s_k$  と  $k$  との関係には「折れ曲がり点」が見出される。たとえば、図 4(a)では、 $m_k, m_k + s_k$  は  $k$  の増大に伴って  $k = 25$  まで減少挙動が見られ、それ以降ではほとんど減少挙動はなく、ほぼ一定の小さな値をとっている。すなわち、このケース（ $t = 12$  時、 $r = 5$ ）では  $k = 25$  において  $m_k, m_k + s_k$  の  $k$  との関係は折れ曲がっている。同様にして、他のケースについても「折れ曲がり点」の  $k$  の値を図から読取る。「折れ曲がり点」の  $k$  の値を、図 3～図 6 から、 $m_k$  の曲線と  $m_k + s_k$  の曲線の両方を併せて判断して求めたのが表 3 である。

表3の結果は表2に示した $k'_i$ と概ね同一である。

表3 折れ曲がり点の $k$ の値

$t_i$	$r = 5$	$r = 1$
10 時	39	39
12 時	25	25
14 時	27	27
16 時	25	25

より直接的に補間分布曲面の $k$ の増大に伴う挙動を検討するため、次の量 $Y_k^s$ を考える。

$$Y_k^s = \frac{\frac{1}{A} \iint_A |P_k^s(x, y) - P_b^s(x, y)| dx dy}{\frac{1}{A} \iint_A |P_k^s(x, y) - [P_k^s]| dx dy} \dots\dots\dots (20)$$

ここで、 $b$ は正の整数で $k \geq b+1$ 。

$Y_k^s$ は、 $b$ を固定して $k$ を動かすとき、基準となる補間分布曲面 $P_b^s(x, y)$ からの $k$ に対応する補間分布曲面 $P_k^s(x, y)$ のへだたりの程度を、 $P_k^s(x, y)$ の $[P_k^s]$ からのへだたりとの比で表わしたものとなっている。

$b = 5$ とし、 $Y_k^s$ を計算した。 $n = 10$ だけある順列のうち、主として次の二つの意味のどちらかで特徴的な順列—— $k$ がある程度大のときには、 $k$ の増大にかかわりなく、 $Y_k^s$ がほぼ一定となる順列、 $k$ が $N$ に近づいてから、大幅に $Y_k^s$ の値が変化している順列——を5個選び出したのが図7～図10であり、これらの図から次のことが言える。濃度分布曲面の形の特徴を支配する測定局を順

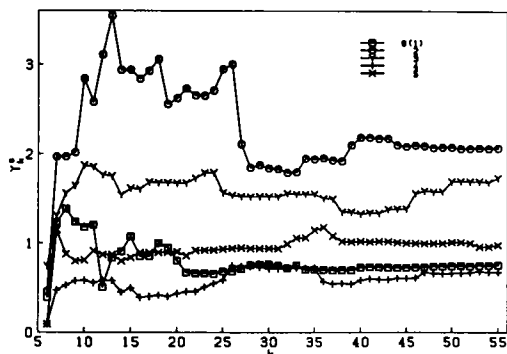


図7  $k$ とそれに伴う補間分布曲面の一定曲面形状への収束の挙動を表わす関数 $Y_k^s$ との関係 ( $t = 10$ 時)

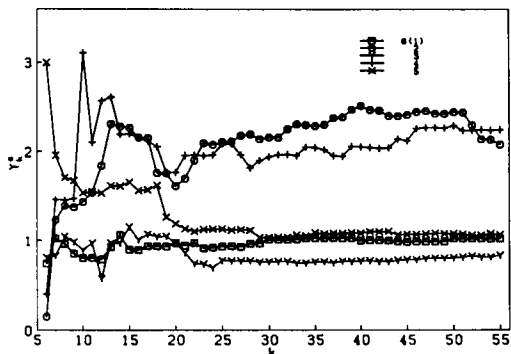


図8  $k$ とそれに伴う補間分布曲面の一定曲面形状への収束の挙動を表わす関数 $Y_k^s$ との関係 ( $t = 12$ 時)

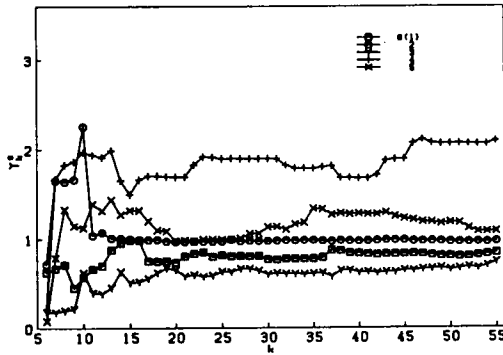


図9 k とそれに伴う補間分布曲面の一定曲面形状への収束の挙動を表わす関数  $Y_k^s$  との関係 (  $t = 1.4$  時 )

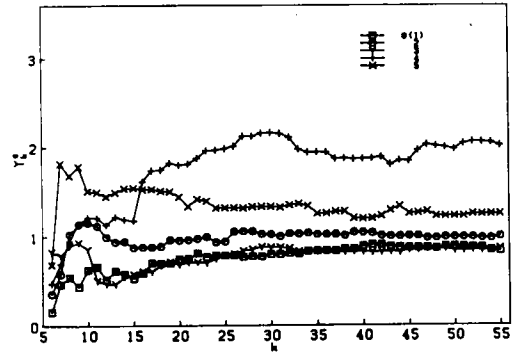


図10 k とそれに伴う補間分布曲面の一定曲面形状への収束の挙動を表わす関数  $Y_k^s$  との関係 (  $t = 1.6$  時 )

列の終り近くで取込まない限り、 $Y_k^s$  はある値に収束してゆく傾向がある。したがって、 $k'$  個程度の測定局のデータで近似曲面が得られていると見なせる。

## 第5節 結 語

離散的な地点で測定された大気汚染データを2次元の分布パターンとして表わすため、データ補間法としてスプライン法の採用を検討した。得られた主な結果は次のようである。

- ① 補間によって作成された分布曲面が真の分布曲面の近似分布曲面と見なせるかどうかを判断する方法を提案した。すなわち、測定局数の増加による補間分布曲面の近似分布曲面への接近の程度を表わす量  $D_k^s / R_k^s$  と、補間分布曲面の一定曲面形状への収束の挙動を表わす量  $Y_k^s$  を定義した。そして  $D_k^s / R_k^s$  が十分長い  $k$  の区間  $[k', k'']$  において十分小であり、 $Y_k^s$  が  $k$  の増加に従って一定の値に収束するならば、その補間方法により近似分布曲面が形成されると判断できることを理論的に示した。
- ② この方法を実際の大気汚染データ（関東地域における55測定局のオキシダントの測定網のデータ）に適用した。まず、補間分布曲面を作成するためのスプライン法のパラメータ  $\sigma$  について検討し、 $\sigma = 1.0 \sim 1.0.0$  (  $h = 1.0 \text{ km}$  ) が適当であるという結果を得た。 $D_k^s / R_k^s$ 、 $Y_k^s$  についての検討結果は次のようであった。測定局数( $k$ )が一定以上に増大するに従い  $D_k^s / R_k^s$  は小さくなり、

一定以上の  $k$  の値では十分小さな値をとりつづけた。また、 $Y_k^s$  は  $k$  の増大に伴って一定値に収束した。この結果により、スプライン法の大気汚染データの補間への適用性が良好であることが確かめられた。

- ③ 以上の手続きにより、近似分布を得るに必要な最小の測定局数が、代表的な分布パターンごとに求められた。この結果を用いて、実際の測定局の最適配置——これについては次章で論ずる——を検討することができる。



## 第7章 大気汚染測定局の最適配置に関する研究<sup>1) 2)</sup>

### 第1節 概 説

環境汚染の測定局の適切な配置のあり方は、モニタリングの目的によって異なる。環境汚染のモニタリングの目的にはいくつかのものがあるが<sup>3)</sup>、大別すると、

- ① 人間や生活環境に対する汚染ばく露の監視
- ② 発生源からの汚染物質の排出の監視
- ③ 汚染対策の評価等のための地域全体の汚染状況の把握

に分類できる。環境汚染の測定局の適切な配置の検討は、これらの目的に応じて行なわねばならない。すなわち、①を目的にした場合には、人口の地域的分布状況等を勘案して、②を目的にした場合では、発生源からの汚染の影響が顕著に現われる場所を対象にして、③を目的にした場合では、地域全体の汚染濃度分布パターン把握に適する場所を対象にして、測定局の設置を考える。しかし、実際の環境汚染モニタリングにおいては、それぞれの目的ごとに別個の測定局が設けられることは少なく、①～③の目的を合わせて満たそうとすることが多い。また、ある一つの目的を満たす測定局配置を検討すると、これが、結果として、他の目的をも満たす測定局であることが多い。たとえば、比較的大きな人口密度や多数の工場を有する都市域において、③を目的とする測定局配置を考えると、①、②の目的も満たすものであることも少なくない。したがって、本章では、③の目的を対象とした測定局配置について論じる。

測定局数が多ければ多いほど一般的には、真の濃度分布に近い濃度分布が得られるが、一方、その設置及び維持管理にそれだけ多くの費用を必要とするので、経済的観点からすれば、測定局の数は少ない方がよい。したがって、測定局の最適配置の問題はできるだけ少ない数の測定局で、できるだけ真の濃度分布に近い分布を推定するためには測定局をどう配置すればよいかを考える問題であると言える。

測定局の最適配置の問題は、現在、例えば、既存都市域における既存の測定局配置をより効率的、合理的なものに見直そうとする場合などにおいて行政的な重

要課題となっている<sup>3)</sup>。

内藤<sup>4)</sup>と落合<sup>5)</sup>とは、大気拡散式を用いて予測した濃度分布を真の濃度分布とみなして、この分布を正しく再現する測定局の配置を検討した。

本章では、かなりの密度で測定局がすでに存在している場合において、その環境汚染の実測データを用いることにより、前章で述べた近似分布と必要な測定局数を求める方法にもとづいて、削減可能局が存在するかどうかを調べ、それが存在する場合に、削減局を選定する方法を提案する。そして、この選定方法を実際のデータに適用し、この方法が有効であることを示す。新規測定局の最適配置は別の問題となる。

## 第2節 測定局の最適削減法

前章によれば、対象地域A上でN個の測定局が既設のとき、そのうちの必要な測定局の数を求めることができる。

前章における定義により、sはN局からN局を選び出す順列のひとつであり、i番目の順列を特定する場合には、s(i)と記す。nは全ての順列のうち、ランダムに選んだ順列の集合、また、その要素数である。P<sub>k</sub><sup>s</sup>(x, y)はsなる順列のうちk番目までの測定局データをスプライン法で補間した曲面である。補間分布曲面P<sub>k</sub><sup>s</sup>(x, y)とこれとの比較曲面P<sub>k</sub><sup>s\*</sup>(x, y)=P<sub>k+r</sub><sup>s</sup>(x, y) r≥1とのへだたりの尺度は、

$$D_k^s = \frac{1}{A} \iint_A |P_k^s(x, y) - P_k^{s*}(x, y)| dx dy \cdots \cdots (1)$$

P<sub>k</sub><sup>s\*</sup>(x, y)のそれ自身の平均値[P<sub>k</sub><sup>s\*</sup>]からのへだたりの程度は、

$$R_k^s = \frac{1}{A} \iint_A |P_k^{s*}(x, y) - [P_k^{s*}]| dx dy \cdots \cdots (2)$$

D<sub>k</sub><sup>s(i)</sup>/R<sub>k</sub><sup>s(i)</sup>の各kでの平均値はm<sub>k</sub>、標準偏差はs<sub>k</sub>と表わされる。

前章によりm<sub>k</sub>, m<sub>k</sub>+s<sub>k</sub>が小でありつづける十分長い区間[k', k'']が存在するとき、n個ある順列のうち、多くのs(i)について濃度分布P<sub>k</sub><sup>s(i)</sup>(x, y), k∈[k', k'']は真の濃度分布P(x, y)に近いと推定され、なおかつ、測定

局の設置数は  $k'$  で不足がないと推定される。すなわち、 $N - k'$  個の局は取除いても、モニタリングの目的は達しうると考えられる。

既設の  $N$  局のうち  $N - k'$  局を削減しても良いとしたときに、どの局を選び出すかを本章で論ずる。

$N$  局から  $N$  局を選び出す順列の全数  $N!$  個の中には、 $k'$  より少ない測定局データで補間の目的を達するものも有ると考えられるが、具体的にそのような順列を探し出そうとすると、 $N!$  個の順列すべてについて総当り的に  $D_k^s / R_k^s$  を求めることになり、多量の計算をしなければならない。これを避けて、本章の考え方では、実用的立場から、 $N - k'$  個のみの削減を行おうとするものである。

$N - k'$  個の測定局を削減するとき、残された  $k'$  個のデータで補間して作った補間曲面と、前章の方法で作った真の分布の近似曲面  $P_a(x, y)$  との差が最小となる順列を用いることが、 $k'$  局で最大の効果をあげるために望ましい。

$N$  個の局のうち  $k'$  個の局を残す組合せのすべてについて、この補間曲面の差を計算することは、事実上不可能である。なぜなら、 $N$  局から  $k'$  局を取出す組合せの個数が通常膨大になるからである。たとえば、 $N = 50$ ， $k' = 20$  とし、その組合せ数は  ${}_{50}C_{20} = 4.7 \times 10^{13}$  となる。そこで、次節のような代替方法を提案する。いくつかの典型的濃度分布パターンデータに対して、それぞれ削減局を求めたのち、すべてのパターンに対して削減してよい局として、パターン毎の削減局に共通している局をとることとする。

### 第3節 削減局決定のための方法

現実的であって、近似的に最適に削減可能局を決定するための方法を以下に与える。

方法は、原理としては、まず、 $N$  個の全測定局の測定データを用いスプライン補間法で分布曲面を作成し、その曲面形から判断して、最も曲面形への影響が小さいとみられる局を取除き、残りの局の値を用いて再び補間分布曲面を作成する。この分布曲面が元の曲面と大差ないことを確認した後、影響の次に小さいと見られる局を取除き、同様の手続きを繰返すものである。

具体的方法の手順は次のとおりである。

- ① まず、補間に用いる測定局の集合  $W_0$  を  $W_0 = \{ N \text{局すべて} \}$  とし、取除きの検討対象となる局の集合  $w_0$  を  $w_0 = W_0$  とする。
  - ②  $W_0$  に対する補間分布曲面を  $P_0(x, y)$  とする。
  - ③  $P_0(x, y)$  曲面の形状から判断して  $w_0$  に属する局のうち曲面形状に与える影響の最も小さいと見られる 1 個の局を  $a$  とし、 $W_1 = W_0 - \{ a \}$ 、また、 $w_0 - \{ a \}$  を  $w_0$  とする。すなわち、 $a$  を取除く。 $W_1$  の測定局のデータの定める補間分布曲面を  $P_1(x, y)$  とする。
  - ④ 
$$Y = \frac{\iint_A |P_0(x, y) - P_1(x, y)| dx dy}{\iint_A |P_0(x, y) - \{P_0\}| dx dy} \dots\dots\dots(3)$$
- $Y$  が十分小さければ⑤へゆく、そうでなければ③へゆく。(このとき、 $a$  を取除くことをやめることになる。以後、 $a$  は取除きの検討対象とならない。)
- ⑤ (取除いた局の総数)  $\geq N - k'$  あるいは、 $w_0$  が空ならば、⑥へゆく。そうでなければ  $W_1$  を  $W_0$  として②へゆく。
  - ⑥ 取除く局を  $\{ N \text{局すべて} \} - W_1$  とする。(手続きを終了する。)

測定局の分布曲面形状へ与える影響の大きさの見積りのために、 $w_0$  の座標  $(x_i, y_i)$  をもつ測定局について次式の値を求め、この値の最小なる局を影響の最小なる局とする。

$$\max \left( \left| \frac{\partial^2 P_0(x_i, y_i)}{\partial x^2} \right|, \left| \frac{\partial^2 P_0(x_i, y_i)}{\partial y^2} \right| \right) \dots\dots\dots(4)$$

(4)式の値の代りに、前章の(1)式で表現される弾性板のエネルギー

$$E = \iint_A [(\Delta Z)^2 + \sigma(\nabla Z)^2] dx dy \dots\dots\dots(5)$$

のうち歪エネルギーに相当する部分  $(\Delta Z)^2$  を用いてもよい。(5)式において  $Z(x, y)$  を  $P_0(x_i, y_i)$  とおきかえて  $(\Delta Z)^2$  を書きかえると

$$\left( \frac{\partial^2 P_0(x_i, y_i)}{\partial x^2} + \frac{\partial^2 P_0(x_i, y_i)}{\partial y^2} \right)^2 \dots\dots\dots(6)$$

(4)式、(6)式のいずれも、分布曲面の曲率の大きさを評価するものであり、どち

らを用いても結果に本質的な差はないと思われる。

前記の方法中で手順④を設けている理由は、次のようである。ある測定局における濃度分布曲面の曲率が小さいからと言って、その測定局の取除きが分布曲面に有意な影響を及ぼさないとは限らない。なぜなら、その測定局の周辺のかかなり広い面積において他の測定局がないような場合などは、当該測定局の取除きは、分布曲面に重大な影響を与えることがあると考えられるからである。

以上の作業により、ある時刻  $t$  のデータの組における分布曲面についての取除き可能測定局が求まる。同様の作業を必要な全ての  $t_i$  における分布曲面に関しても行い、それぞれの取除き可能測定局から共通の測定局を選び出せば、これが求めるものである。

選び出す必要のある時刻  $t_i$  は代表的濃度分布を示す時刻  $t_i$  である。具体的には、十分長い時間（たとえば1年間）の中で濃度測定がされた時刻  $t$  の全て（たとえば  $24/\text{日} \times 365\text{日}$ ）のうちから、高い濃度が出現した時刻を取出す。この選び出された時刻を  $t_1, t_2, \dots, t_q$  とする。汚染濃度が低い場合を選ばないのは、測定の意義が小さいこと、それにもかかわらず、濃度曲面の起伏が小さいため、領域全体での平均濃度からのへだたり  $R$  が小であるので、かえって測定局を多く必要とする結果を与えるからである。

#### 第4節 実データへの適用

実際の測定データに第3節の測定局取除きの手続きを適用する。用いたデータおよび対象領域は前章と同じである。選び出した時刻  $t_i$  は高い濃度が出現した時刻  $t_1 = 10\text{時}$ 、 $t_2 = 12\text{時}$ 、 $t_3 = 14\text{時}$ 、 $t_4 = 16\text{時}$  である。

測定局の分布曲面形状への影響の見積りには(4)式を用いた。手順のステップ④の  $Y$  の大きさの判断基準を  $0.05$  とした。

時刻  $t_i$  における分布曲面について、手順を実行して、取除かれた順番に測定局に1から順に番号をつける。この番号を  $J_i$  とすると、 $J_i \leq N - k'_i$  である番号で指定される測定局は取除き可能局である。その個数は  $N - k'_i$  個、取除いた後の測定局の数は  $k'_i$  である。用いたデータについての  $k'_i$  は前章4.2の議論に

より、表 1 のように求められている。表 2

表 1  $k'_i$  の値

では、各  $t_i$  のそれぞれについての取除き可能局を示した。番号がついている局が  $t_i$  における取除き可能局である。全ての  $t_i$  について番号が付されている局を選び出せば、求める共通な 5 個の取除き可能局（以下、これを削減可能局と称す。）が求められる。これを○印で示した。

$t_i$	$k'_i$
$t_1$ (10時)	37
$t_2$ (12時)	25
$t_3$ (14時)	27
$t_4$ (16時)	25

表 2 各測定局の  $J_i$  と削減可能局（○，△印の意味は本文第 4 節参照）

局 コード	$J_i$				共通削減 可能局	局 コード	$J_i$				共通削減 可能局	局 コード	$J_i$				共通削減 可能局
	$t_1$	$t_2$	$t_3$	$t_4$			$t_1$	$t_2$	$t_3$	$t_4$			$t_1$	$t_2$	$t_3$	$t_4$	
1						21	10		13			41	5		25	28	
2						22				19		42			24	27	
3		28	16			23	15	26	11	4	○	43			21		
4		11				24	8	18	20			44	16	24		20	
5		4	26			25						45					
6	1	27				26	18			17		46				25	
7		30	18			27	2	19	22	9	○	47	13		23	16	
8	11	8	27			28	17	5	28	21	○	48					
9		6	3	10	△	29	12	15		30		49					
10		20	6	23	△	30		13		26		50			14	8	
11		12				31	7	1	7	11	○	51				24	
12	14	21		29		32			8	12		52			17	14	
13	4	22		7		33		23				53	7		4	1	△
14		2	12			34				6		54			1	3	
15			10	18		35		16	5			55		9		2	
16		10				36	3	3				削減可能 局 数 ( $N-k'$ )					
17				15		37		14	2	13	△		18	30	28	30	5
18		25	9			38		17	15	5	△						(10)
19						39											
20	9	29	19	22	○	40	6										

各  $t_i$  は高い汚染濃度が出現する時刻として選び出されたものであるが、このうち、 $t_1 = 10$  時の場合は、光化学スモッグの発生が始まったばかりの段階であり、濃度分布曲面の起伏が他と比べてさほど著しくない場合であるので、第 3 節で述べた理由により、取除き可能局を検討する対象から除外することも考えられる。その場合、前述の 5 局（○印を付したもの）に新たに追加される 5 個の削減可能局（表 2 中△印を付したもの）の合計 10 局である。

図 1 に、削減可能局と残りの測定局の位置関係を示した。削減可能局については、前述と同じ意味の○印と△印で囲んで示した。

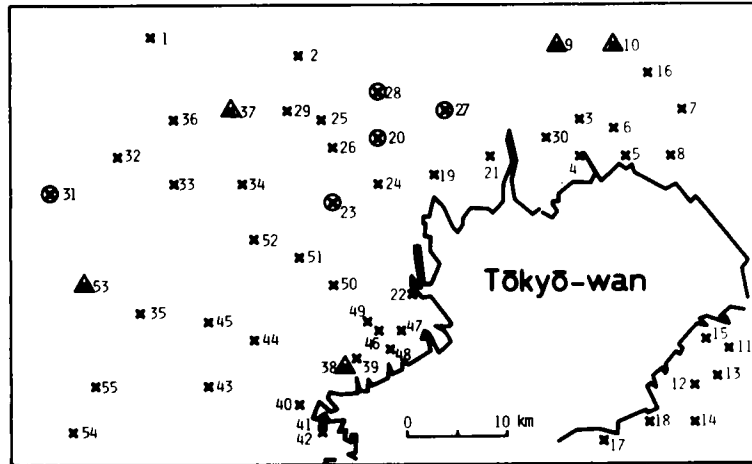


図1 対象領域内の測定局と削減可能局（○，△印の意味は本文第4節参照）

## 第5節 結果の検討

これまでに述べて来たところにより、取除いても結果（分布曲面）に大きな差をもたらさない測定局（削減可能局）が明らかとなった。このような測定局の有する条件としては、濃度ピークからはずれた地域で、濃度分布の2次微分の小さい地域、または測定局の配置密度が高い地域に存在するものであることが想定される。図2(a)～図5(a)は、それぞれ  $t_1 \sim t_4$  について、全測定局のデータを用いて補間した近似分布曲面  $P_a(x, y)$  である。これらの図と対照すれば、図1において、求められた10個の削減可能局は、2次微分の大きい濃度分布曲面のピークから外れていることが分る。

削減可能局が正しく指定されているか否かを検討するため、削減可能局として指定された局のデータを用いなくて補間曲面を作成し、これを、全ての局（55局）のデータを用いて作成した補間曲面と比較して、差がないかどうかを調べる。図2(b)～図5(b)は、それぞれ  $t_1 \sim t_4$  について、5個の削減可能局（表1で○印を付けた局）のデータを用いなくて補間した分布曲面である。視覚的に、両者はほとんど同一の分布曲面に見える。すなわち、濃度の高い部分、低い部分は、それぞれ互いにはほぼ一致し、濃度等高線の概形も一致している。また、図2(c)～図5(c)は、それぞれ  $t_1 \sim t_4$  における取除き可能局（表2中  $J_i$  の番号がつ

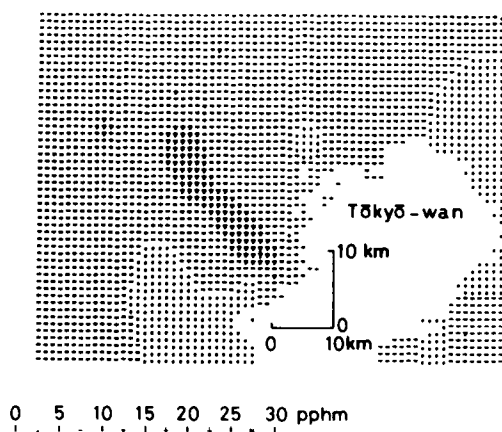


図 2 (a) 補間分布曲面 (  $t=10$  時,  $k=55$  )

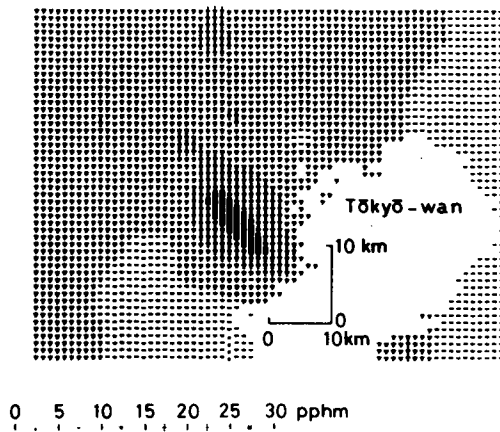


図 3 (a) 補間分布曲面 (  $t=12$  時,  $k=55$  )

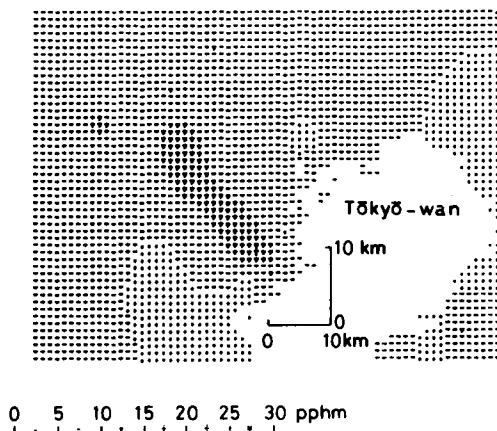


図 2 (b) 補間分布曲面 (  $t=10$  時,  $k=50$  )

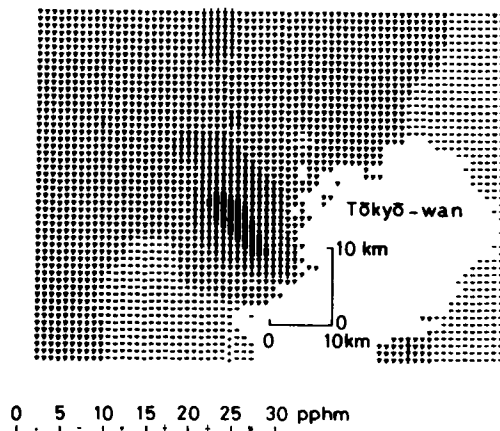


図 3 (b) 補間分布曲面 (  $t=12$  時,  $k=50$  )

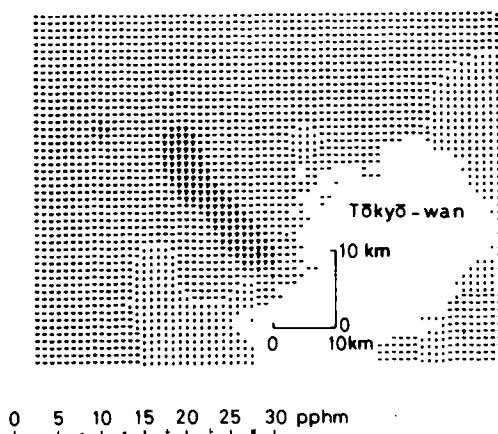


図 2 (c) 補間分布曲面 (  $t=10$  時,  $k=37$  )

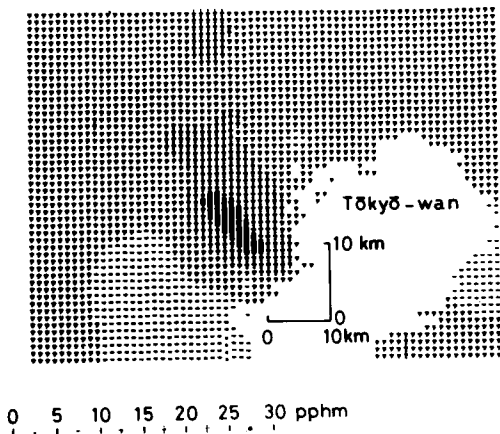


図 3 (c) 補間分布曲面 (  $t=12$  時,  $k=25$  )



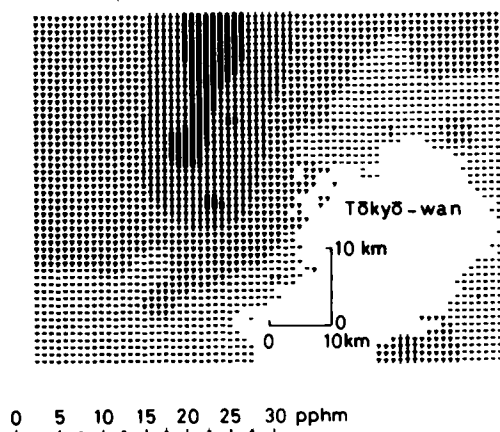


図 4 (a) 補間分布曲面 (  $t=14$  時,  $k=55$  )

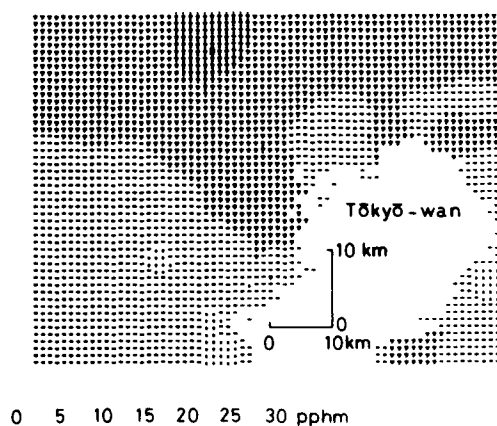


図 5 (a) 補間分布曲面 (  $t=16$  時,  $k=55$  )

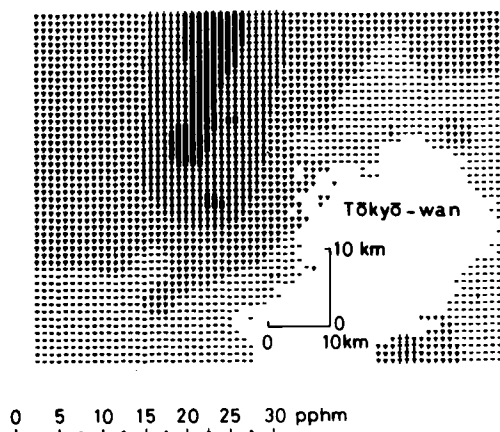


図 4 (b) 補間分布曲面 (  $t=14$  時,  $k=50$  )

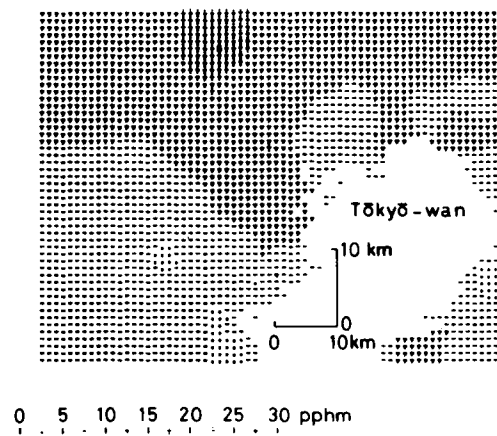


図 5 (b) 補間分布曲面 (  $t=16$  時,  $k=50$  )

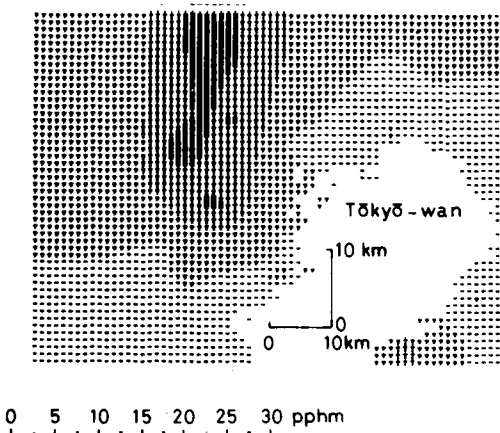


図 4 (c) 補間分布曲面 (  $t=14$  時,  $k=27$  )

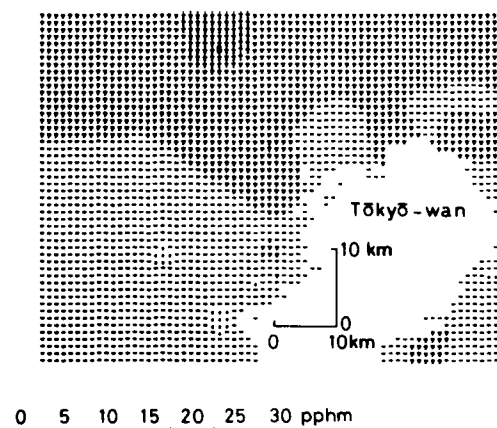


図 5 (c) 補間分布曲面 (  $t=16$  時,  $k=25$  )

けられているそれぞれ 18, 30, 28, 30 個の測定局)のデータを用いないで補間した分布曲面である。図 2 (a)～図 5 (a)の対応する分布曲面と対照しても上記の意味での視覚的な差は、図 2 (b)～図 5 (b)のときよりも、やや大きいだけである。

次に、削減可能局決定のためにデータを用いていない時刻(すなわち  $t_1 \sim t_4$  以外の時刻)についても調べる。同日の  $t_1' = 11$  時、 $t_2' = 13$  時、 $t_3' = 15$  時、 $t_4' = 17$  時について、全測定局のデータを用いて補間した分布曲面、図 6 (a)～図 9 (a)と、削減可能局(表 2 で○印を付けられた 5 局)のデータを用いないで補間した分布曲面図 6 (b)～図 9 (b)を比較する。この場合も両者に大きな差は認められない。

よって、削減可能局が正しく指定されていることが検証された。そして、このことにより、本章で論じた方法による測定局の削減方法は、濃度分布パターンの適切な把握という観点から考えた場合、十分採用しうる方法であることが実証された。

なお、本章における測定局の最適配置の考え方は、濃度分布パターンを適切に把握するのに必要な測定局配置という観点からの一つの考え方に過ぎない。一般論として測定局の最適配置を考えるには次の点の検討が必要である。すなわち、モニタリングの目的(第 1 節で論じた)を検討し、それに応じた最適配置を考え

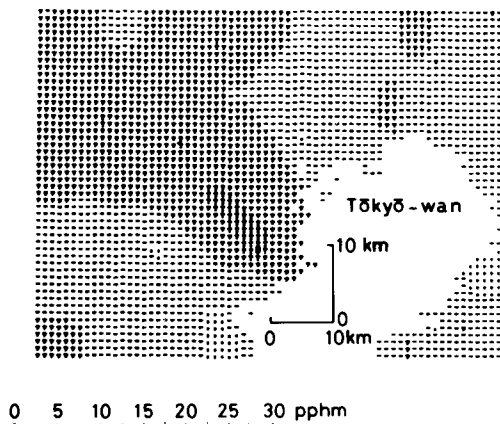


図 6 (a) 補間分布曲面 ( $t=11$  時,  $k=55$ )

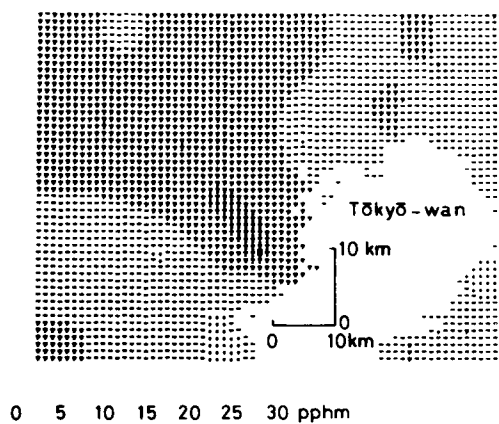


図 6 (b) 補間分布曲面 ( $t=11$  時,  $k=50$ )

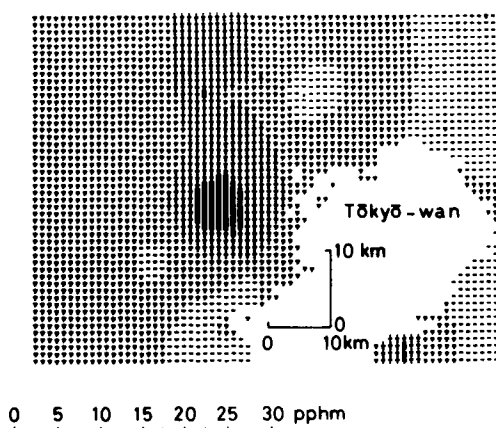


図 7 (a) 補間分布曲面 (  $t=13$  時,  $k=55$  )

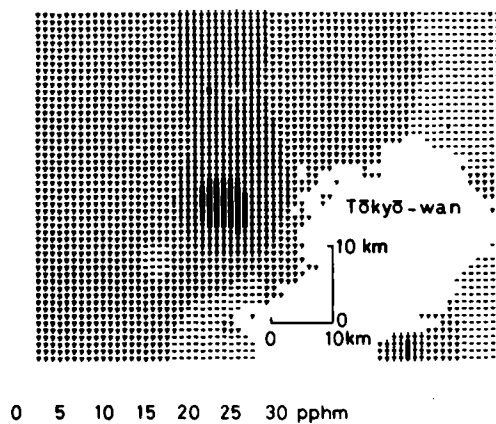


図 7 (b) 補間分布曲面 (  $t=13$  時,  $k=50$  )

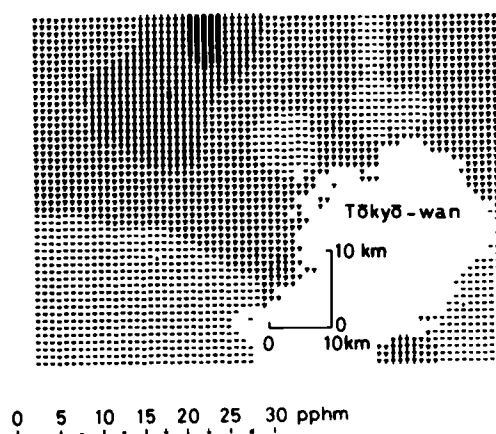


図 8 (a) 補間分布曲面 (  $t=15$  時,  $k=55$  )

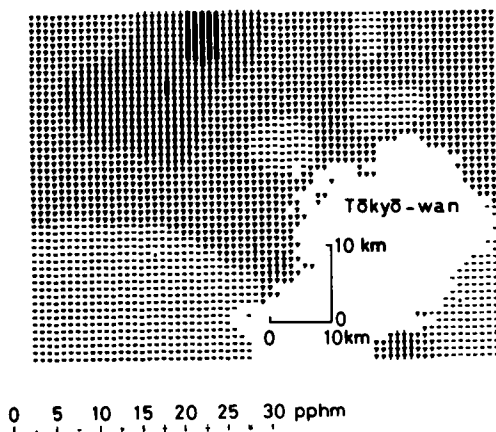


図 8 (b) 補間分布曲面 (  $t=15$  時,  $k=50$  )

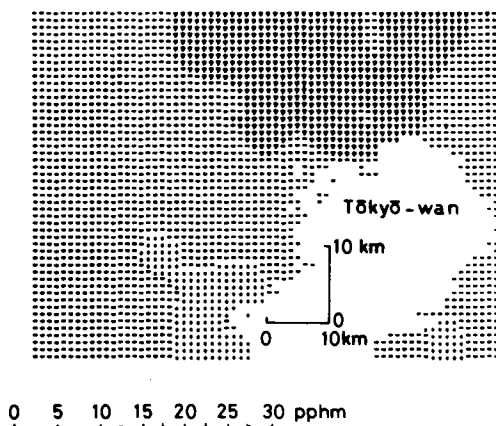


図 9 (a) 補間分布曲面 (  $t=17$  時,  $k=55$  )

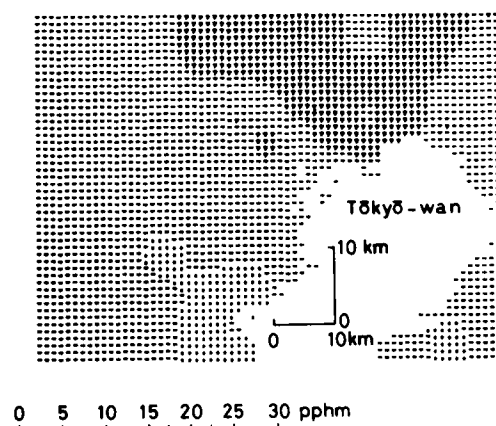


図 9 (b) 補間分布曲面 (  $t=17$  時,  $k=50$  )

る必要があること、対象とする測定局の実際上の設置場所は測定局の維持管理のしやすさ等実務上の観点からの検討も必要であることなどである。したがって、本節において削減可能局として選定された測定局が実際のモニタリングシステムにおいて削減すべき測定局であるとするには、上記の検討が追加されねばならない。

## 第 6 節 結 語

既存の大気汚染測定網から取除くことが可能である測定局を選定する方法を提案した。この方法は、前章で論じた必要な測定局の数 ( $N - k'$ ) を求める方法に基づき、代表的な汚染濃度分布について、取除き可能な局を準最適に選定するものである。方法の概要は次のようである。

- ① 濃度分布の曲面形から判断して、最も曲面形への影響が小さいとみられる 1 局を取除く。(曲面形への影響は(4)式の値に基づいて判断する。)
- ② 残りの局の値を用いて補間分布曲面を作成し、この分布曲面が元の分布曲面と大差ないことを確認する。(差の程度の判断は(3)式によって行う。)
- ③ 曲面形への影響が次に小さいとみられる 1 局を同様にして取除き、同様の操作を、取除く局の総数が  $N - k'$  になるまで続ける。
- ④ 代表的な汚染濃度分布の全てについてこの操作を行い、それぞれにおいて求められた取除き可能局の組から共通な局を抽出すれば、求める削減可能局が得られる。

この方法を実際の大気汚染測定網に適用し削減可能局を選定した。関東地域における 55 測定局のオキシダントの測定網の例では、5 局が削減可能局として選定された。削減可能局のデータを用いずに補間した分布曲面と全測定局のデータを用いて補間した分布曲面とを比較し、両者の分布曲面の形状に大きな差がないことを確認した。そしてこのことにより、本章で述べた削減可能局の選定法は実用的に採用しうる方法であることを実証した。

## 第8章 水質環境データの2次元化と合理的 測定点配置に関する研究<sup>1)</sup>

### 第1節 概 説

湖や海における水質汚濁の状況を全体的に把握、評価するためには、水域全体の水質汚濁状況を二次元的な濃度分布パターンとして把握することも有効な1つの方法である。本章ではスプライン法によるデータ補間法を湖の水質濃度分布パターンの推定に適用し、その実用性を実証する。まず、同法を水質汚濁データに適用する際の収束条件、パラメータの与え方、計算の迅速化のための技術などを検討する。そしてこのデータ補間法を琵琶湖南湖の実データに適用し、その水質濃度分布パターン及びその季節的変化が視覚的に適切に把握されることを示す。ついで、この濃度分布推定法を用いて測定点の合理的な配置を決定する方法を検討する。前章では一つの汚染物質（オキシダント）についての測定局の配置問題について検討したが、ここでは複数の項目を測定対象としている測定点（水質汚濁の測定点はこのような場合が多い。）の配置問題を検討する。このため、分布曲面に与える測定点の相対的な重要性を評価することにより、削減可能測定点及び代表測定点を選定する方法を提案する。また、この方法を実際のデータに適用することによりその実用性を確める。検討対象としたデータは、昭和53年1月から12月まで、1か月に2回ずつ、13水質項目（クロロフィルa（以下Chl.aと略記する）、全COD（重クロム酸法）（T-COD）、浮遊性COD（P-COD）、溶解性COD（S-COD）、全窒素（T-N）、有機態窒素（Org-N）、無機態窒素（In-N）、硝酸性窒素（NO<sub>3</sub>-N）、アンモニア性窒素（NH<sub>4</sub>-N）、全リン（T-P）、リン酸態リン（PO<sub>4</sub>-P）、浮遊性リン（P-P））について、22測定点で測定された琵琶湖南湖の水質測定結果<sup>2)</sup>である。なお、採水水深は表面下50cmである。

## 第2節 スプライン法の水質汚濁データへの適用

### 2.1 スプライン法と水質汚濁データへの適用上の課題

スプライン法の定式化と差分方程式へのおきかえについては第6章で述べた。

離散的な数の限られた地点について得られた水質汚濁データから連続的な二次元濃度分布を推定するためにスプライン法を用いる場合、とくに次のような点の検討が重要である。まず、逐次近似計算を行う場合の計算の収束性に関する検討が必要である。すなわち、水質汚濁の測定のようにデータ点が少ない場合はスプライン補間の収束性は非常に悪い場合が多いので、これの向上を図るための措置を講ずること、またこれの前提として計算の収束をどのように判定するか——収束条件の設定——を検討することが必要である。次に、補間濃度分布ができるだけ正しく推定される（実測値と補間によって推定された値との差を小さくする）よう、パラメータの最適値域を設定することが必要である。

### 2.2 対象領域、境界条件、初期値の設定

二次元水平面（ $x, y$ ）上のデータ点におけるデータの値（データ数： $I$ ）に相当する高さを $z$ とし、測定点 $(x, y) = (x_i, y_i)$ , ( $i = 1, 2, \dots, I$ )において、それぞれ $z = z_i$ , ( $i = 1, 2, \dots, I$ )となるようなスプライン補間曲面 $z = Z(x, y)$ を求めることを考える。計算の範囲を琵琶湖南湖の水面の領域を含む長方形（縦 $1.7\text{ km}$ 、横 $6.5\text{ km}$ ）の領域とした（図1参照）。メッシュ幅（代表領域の幅）を $h = 0.25\text{ km}$ とし、格子数は縦 $q = 6.8$ 、横 $p = 2.6$ とした。代表領域（以下メッシュと呼ぶ。）に1つのデ

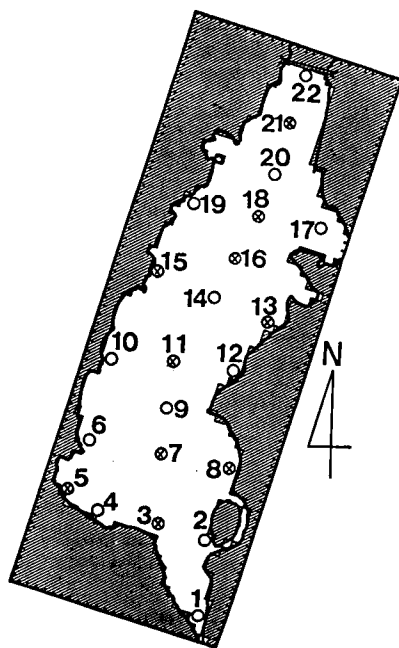


図1 測定点の位置と番号

ータが存在するとき、それをメッシュに与えるデータとして採用した。

境界条件として、境界における1次、2次の微分係数を0とした。つまりメッシュ  $r_{\alpha, \beta}$  の  $Z$  を  $Z_{\alpha, \beta}$  と表わすと、 $Z_{1, \beta} = Z_{3, \beta}$   $Z_{2, \beta} = Z_{3, \beta}$  (ただし、ここで  $\beta = 3, 4, \dots, q-2$ ) である。 $Z_{\alpha, 1}$  ,  $Z_{\alpha, 2}$  についても同様である。

スプライン法による計算を開始するとき、データ点が存在しない格子点に次の方法により初期値  $Z_{\alpha, \beta}^0$  を与えた。

$$Z_{\alpha, \beta}^0 = \{ (z_1)_1 \cdot d_2 + (z_1)_2 \cdot d_1 \} / (d_1 + d_2) \quad \dots\dots\dots(1)$$

(ただし、 $(z_1)_1, (z_1)_2$  は格子点  $(X_\alpha, Y_\beta)$  に最も近い2つのデータの値、 $d_1, d_2$  は格子点からそれぞれのデータ点までの距離の2乗である。) 一方、データの存在する格子点ではその実測値を初期値として与えた。

### 2.3 収束条件の検討

逐次近似法の計算は繰返し近似計算を行うことにより最終解に到達する。その際どの程度の回数で計算を打ち切るかの判断をする必要がある。ここでは、 $J$  回目と  $(J+1)$  回目の解の差が一定の値以下になるときをもって計算を打ち切る方法を採用した。 $J$  回目、 $(J+1)$  回目の解の差  $\epsilon_m, \epsilon_a$  を次式によって与え、それぞれ逐次誤差最大値、逐次誤差平均値と呼ぶ。これらがどの程度の値より小さくなれば、実用上収束したとみなすか検討する。

$$\epsilon_m = \text{Max}_{(X_\alpha, Y_\beta) \in A'} |Z_{\alpha, \beta}^J - Z_{\alpha, \beta}^{J+1}| / a \quad \dots\dots\dots(2)$$

$$\epsilon_a = \sum_{(X_\alpha, Y_\beta) \in A'} |Z_{\alpha, \beta}^J - Z_{\alpha, \beta}^{J+1}| / M \cdot a \quad \dots\dots\dots(3)$$

ただし、 $Z_{\alpha, \beta}^J, Z_{\alpha, \beta}^{J+1}$  はそれぞれ  $J, J+1$  回目における  $Z_{\alpha, \beta}$  の計算値、 $A'$  は計算対象領域、 $M$  は格子点数、 $a$  は測定データの平均値である。

つぎに、 $J$  回目の解と最終回の解との差  $\eta_m, \eta_a$  を次式によって与え、それぞれ最大誤差、平均誤差と呼ぶ。

$$\eta_m = \max_{(X_\alpha, Y_\beta) \in A'} |Z_{\alpha, \beta}^J - Z_{\alpha, \beta}^\infty| / a \quad \dots\dots\dots(4)$$

$$\eta_a = \sum_{(X_\alpha, Y_\beta) \in A'} |Z_{\alpha, \beta}^J - Z_{\alpha, \beta}^\infty| / M \cdot a \quad \dots\dots\dots(5)$$

ただし、 $Z_{\alpha, \beta}^\infty$  は  $J = \infty$  のときの  $Z_{\alpha, \beta}$  の計算値である。ここでは実際の計算結果ではばまるめの誤差内に収束したとみなせる  $J = 2000$  回をもって  $J = \infty$  と見なすこととした。

最大誤差または平均誤差を実用上問題ない程度に小さくするに必要な逐次誤差最大値および逐次誤差平均値についての条件を、実データを適用することにより調べる。

ここでは、スプライン・パラメータ  $\sigma h^2$  を 1.0 に設定した。図 2 は最大誤差 ( $\eta_m$ )、平均誤差 ( $\eta_a$ ) と繰り返し計算回数 ( $J$ ) との関係を示したものである。検討対象に用いたのは昭和 53 年測定の Chl. a の年平均値である (表 1 参照)。図 2 より  $\eta_m$  の収束値は約  $3 \times 10^{-2}$ ,  $\eta_a$  の収束値は約  $2 \times 10^{-3}$  であることがわかる。 $\eta_m$ ,  $\eta_a$  をこの程度の値におさめれば、計算値と最終解との差は格子点の値の差の絶対値の最大で約 3%, 平均で 0.2% となる。この程度の

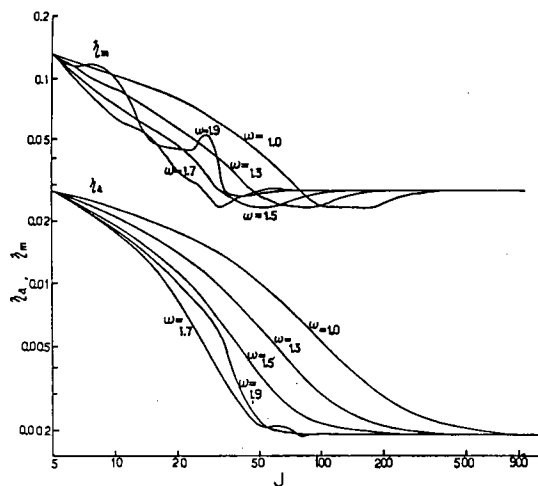


図 2 最大誤差、平均誤差と繰り返し計算回数との関係

差は、水質の測定値と測定限界の関係——現在よく用いられている Chl. a の測定では測定値の 95% 信頼限界は「平均値 ± (平均値の約 5%)」<sup>3)</sup>といわれる——から実用上十分許容されるものである。



表 1 琵琶湖南湖の水質測定値（昭和 53 年度平均値）

Unit: mg/l (\*), µg/l (no mark)

Sta. No.	Chl. a	SS	T-COD	P-COD	S-COD	T-N	Org-N	In-N	NO <sub>3</sub> -N	NH <sub>4</sub> -N	T-P	PO <sub>4</sub> -P	P-P
1	15.7	7.3	8.2	3.1	5.1	669	567	103	86	18	44	3	40
2	15.5	6.7	7.8	2.7	5.1	556	460	96	71	25	42	6	36
3	12.2	5.5	6.9	2.0	5.2	456	356	100	73	27	37	3	34
4	12.1	6.3	7.3	2.3	5.2	442	353	88	71	18	35	3	32
5	12.3	5.2	7.0	2.1	4.9	423	340	83	66	17	39	4	35
6	14.3	5.9	7.0	1.8	5.1	553	475	79	64	15	58	4	54
7	12.6	5.4	6.9	2.0	4.9	418	342	76	63	13	37	3	34
8	16.2	6.8	7.8	2.5	5.4	517	424	93	71	22	49	4	45
9	13.3	5.4	6.7	2.0	4.9	393	310	82	69	14	43	3	40
10	13.1	4.7	6.8	2.0	4.8	398	317	81	67	14	35	3	32
11	12.5	5.9	6.5	1.9	4.6	366	279	86	68	19	33	2	31
12	22.1	9.5	8.6	3.3	5.4	552	378	174	97	78	63	6	57
13	15.6	9.5	7.2	2.6	4.6	442	330	112	87	25	55	4	51
14	9.9	5.0	6.2	1.9	4.6	351	259	92	73	19	29	3	26
15	10.8	4.1	6.4	2.3	4.1	355	268	88	71	17	33	2	31
16	10.5	8.1	6.0	1.5	4.5	463	373	90	69	21	35	3	32
17	14.4	9.4	7.1	2.4	4.7	452	326	126	87	39	51	3	48
18	9.6	4.3	6.4	1.4	5.2	349	264	85	68	17	32	2	30
19	9.9	4.5	5.8	1.6	4.2	351	260	91	73	18	33	3	31
20	9.9	5.3	5.7	1.5	4.2	345	260	85	67	17	39	3	37
21	10.2	4.3	5.3	1.4	4.1	332	235	97	77	20	30	3	27
22	7.9	2.7	5.0	1.2	3.9	300	216	84	70	14	23	1	22
AVR.	12.75	5.99	6.75	2.08	4.75	431.0	336.0	95.0	73.0	22.1	39.8	3.3	36.5
S.D.	3.04	1.79	0.89	0.55	0.42	90.0	84.2	20.5	8.4	13.4	9.7	1.0	9.0

最大誤差 ( $\eta_m$ )、平均誤差 ( $\eta_a$ ) をこの程度の値におさめるためには逐次誤差最大値 ( $\varepsilon_m$ )、逐次誤差平均値 ( $\varepsilon_a$ ) をそれぞれどの程度の値におさめればよいかを見る。計算繰り返し回数  $J$  の増加による  $\eta_m$ ,  $\eta_a$ ,  $\varepsilon_m$ ,  $\varepsilon_a$  の減少の状況を  $\omega$  (次項で検討) が 1.9 の場合について図 3 に示した。図から  $\eta_m$ ,  $\eta_a$  が収束値に到達するときの  $\varepsilon_m$ ,  $\varepsilon_a$  はそれぞれ  $10^{-4}$ ,  $10^{-5}$  のオーダーであることがわかる。他の  $\omega$  の値のときについても同様のことが言える。したがって、収束条件として、 $\varepsilon_m < 10^{-4}$ ,  $\varepsilon_a < 10^{-5}$  を用いれば実用的

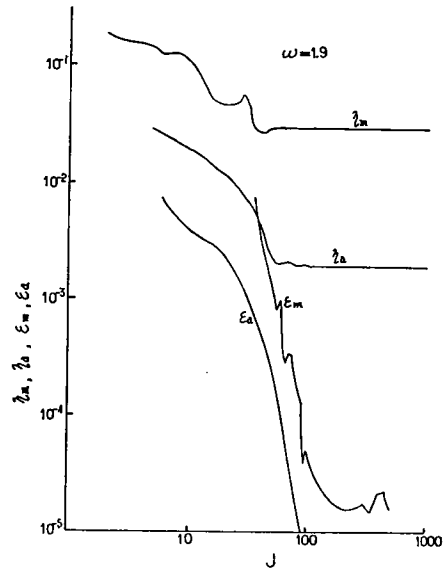


図 3 繰り返し計算回数による  $\eta_m$ ,  $\eta_a$ ,  $\varepsilon_m$ ,  $\varepsilon_a$  の変化

に十分な精度の結果が得られると言える。なお、この条件に加え、図 3 から判断して、繰り返し回数 200 回を計算打ち切り条件として採用した。

## 2.4 収束性の改良

逐次近似法による連立方程式の解法（ガウス・ザイデル法）は収束性が問題である。とくに、測定点が比較的少なく、メッシュが細かい——水質データのスプライン補間もこの例であることが多い——場合は収束の速度は非常に遅くなる。収束速度の向上を図るため SOR (Successive Over-relaxation) 法<sup>4)</sup>の適用を試みる。この方法では、第  $k$  近似解  $x^{(k)}$  に修正量  $\Delta x^{(k+1)}$  を  $\omega$  倍したものを加え  $x^{(k+1)}$  を作る方法である。 $\omega$  は加速パラメータと呼ばれており<sup>4)</sup>、その存在域は  $0 < \omega < 2$  である。平面応力解析の問題には  $\omega$  の値は 1.8 から 1.95 の間がよいと言われている<sup>4)</sup>。

$\omega = 1.0, 1.3, 1.5, 1.7, 1.85, 1.9, 1.95$  について、 $J$  の増加に伴う  $\eta_m, \eta_a$  の減少状況を調べた。結果の概要を図 2 に示す。用いたデータは前述と同じ（測定点数は 22）である。 $\omega$  が 1 から 1.7 までの間ではその値が大きくなるほど  $\eta_m, \eta_a$  の収束速度は大きくなるが、それ以上 1.9 までは同程度かやや小さくなる傾向がある。この関係を図 4 に示した。図から、測定点が比較的多い場合は  $\omega$  は 1.7 ～ 1.9 程度が適当であることがわかる。

次に、データ点が極端に少ない場合——例えば、データ点が 1 点しか存在しないときに補間分布を作成するという特殊な場合——の収束状況について検討する。データ点が 1 点のみの場合の補間分布はその値の様な平面分布となるが、初期値の与え方によっては収束がきわめて悪いのが問題である。図 5 は、22 測定点の

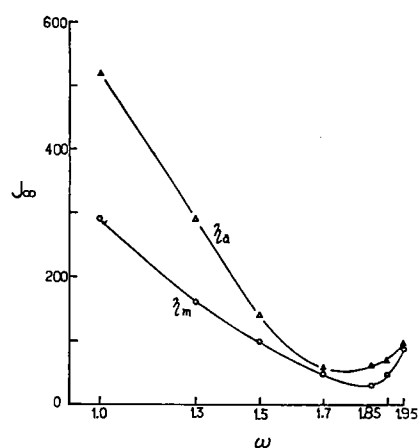


図 4 収束に至るまでの繰り返し計算回数 ( $J_{\infty}$ ) に与える  $\omega$  の影響

データを用いて計算した結果を初期値として与え、 $\omega$ の値ごとに収束状況を調べたものである。収束性は図2で得られたそれと比較してきわめて悪い。しかし、 $\omega = 1.0$  から  $\omega = 1.95$  まで  $\omega$  の値が大きくなるほど収束性がよくなることがわかる。

このことと前述のことから判断して、実用上  $\omega = 1.9$  付近が適当であると考えられる。以下の計算では  $\omega = 1.9$  を用いる。図4より、加速パラメータ  $\omega = 1.9$  を用いたSOR法は、単純なガウス・ザイデル法 ( $\omega = 1.0$  の場合に相当) より収束に至る繰り返し回数を1オーダー程度短縮することができる。  $\sigma$  の値によっても収束性を向上させることができるがこの点の検討は次項で行う。

## 2.5 パラメータ $\sigma$ の値の検討

スプライン法を用いる場合は、唯一のパラメータである  $\sigma$  の値を適切に設定することが重要である。大気汚染データについてのパラメータ  $\sigma$  の最適値の検討は前章で論じたが、本章では水質汚染データについてのそれについて調べる。用い

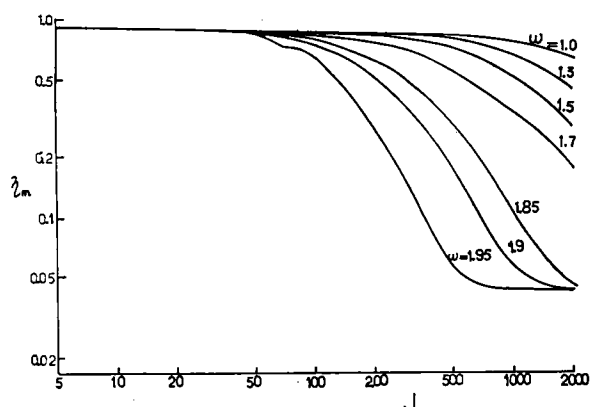


図 5(a) データ点が1点の場合の収束状況

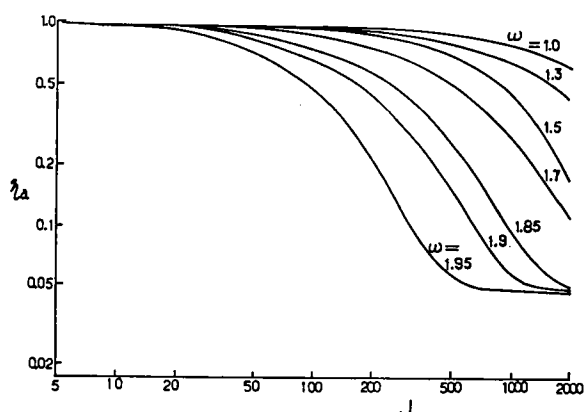


図 5(b) データ点が1点の場合の収束状況

るデータは前述の琵琶湖南湖の水質データ（表 1 参照）で、用いる方法は次のようである。琵琶湖南湖に存在する 22 個の測定点（測定点の位置と番号は図 1 に示した）のうち 10 測定点（同図に⊗で示した）の測定値を用い、補間分布曲面を作る。測定値を用いなかった測定点  $i$  の測定値を  $z_i$ 、補間によって得られる測定点  $i$  における推定値を  $\hat{z}_i$  とした場合の

$$y = \frac{\sum_{i=1}^{10} (z_i - \hat{z}_i)^2}{\sum_{i=1}^{10} (z_i - \bar{z})^2}, \text{ただし } \bar{z} = \sum_{i=1}^{10} z_i / 10$$

の値を、いくつかの  $\sigma h^2$  の値について求め、これが小となる  $\sigma h^2$  を求める。各水質項目について  $\sigma h^2$  と  $y$  の関係を調べた結果を図 6 に示した。図から、 $\text{NH}_4\text{-N}$  を除いて、いずれの水質項目についても、 $\sigma h^2$  の値を変化させても  $y$  の値はそれほど大きく変化しないことがわかる。

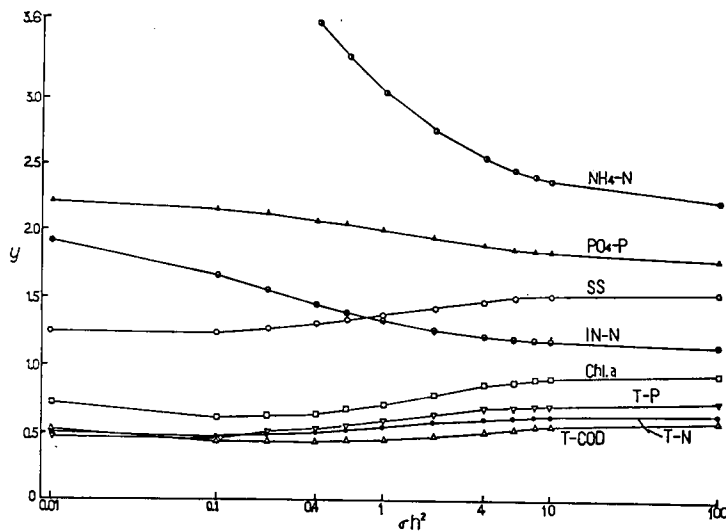


図 6  $\sigma h^2$  と誤差( $y$ )との関係

一方、 $\sigma h^2$  の値の収束性への影響をみるため、Chl.a の年平均データを用い、 $\sigma h^2$  と収束に至る繰り返し計算回数 ( $J_\infty$ ) の関係を調べた（図 7）。図から、 $\sigma h^2$  が大きくなるほど  $J_\infty$  が小さくなる、すなわち収束性がよくなることがわかる。例えば  $\sigma h^2 = 1.0$  のときは  $\sigma h^2 = 0.1$  のときより収束に至る繰り返し計算回

数を  $1/2$  以下に短縮することができる。

この観点から  $\sigma h^2$  は余り小さくない方が望ましい。また、補間分布曲面の滑らかさを保つためには極端に大きな  $\sigma h^2$  は避ける方が望ましい<sup>5)6)</sup>。大気汚染データのスプライン補間で得られた  $\sigma h^2$  の最適値域は前章で論じたとおり  $1 \sim 10$  であったが、以上の考察を踏まえ、水質汚濁データの場合、 $\sigma h^2$  の値はこの範囲で十分であるということが知られる。本章において、以下の計算では  $\sigma h^2 = 1.0$  を採用することとした。

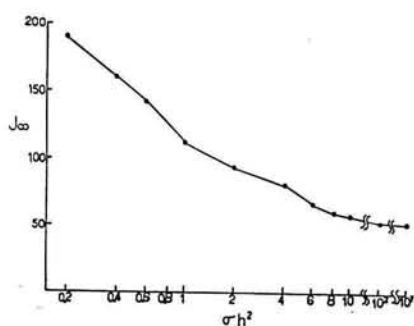


図7 収束に至るまでの繰り返し計算回数 ( $J_{\infty}$ ) に与える  $\sigma h^2$  の影響

### 第3節 スプライン法による琵琶湖南湖の水質汚濁濃度分布パターンの推定

前述の検討結果を用い、琵琶湖南湖の水質汚濁濃度の分布パターンとその季節的变化を推定する。昭和53年の1月～12月、1か月に2回ずつ測定された Chl.a のデータを用い、それぞれの濃度分布パターンを推定した。そのうちの一部を図8に示した。このように Chl.a 濃度を二次元濃度分布パターンで表現することにより、濃度は特定の時期、特定の地点で高くなること、高い濃度が出現す

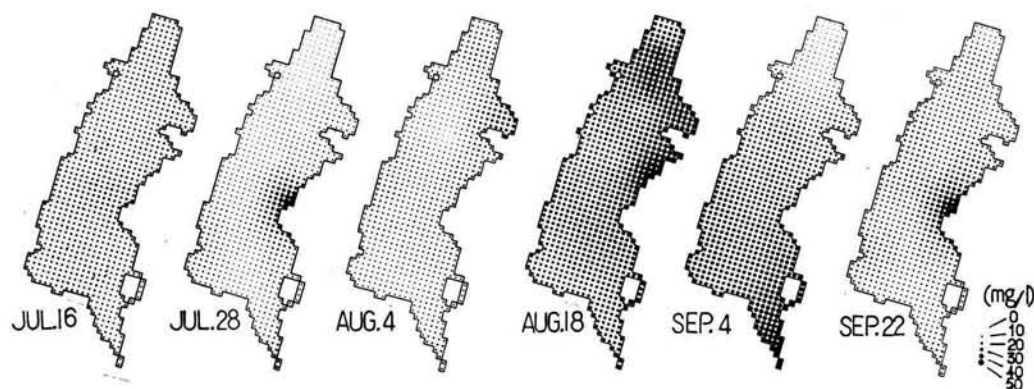


図8 スプライン補間法による Chl.a の濃度分布パターン

る測定時の濃度分布（例えば8月18日の濃度分布）を下流方向へ押し下げたような形で次の測定時の濃度分布（9月4日の濃度分布）が出現していることなどが視覚的に容易に把握しうる。

## 第4節 合理的測定点配置の検討

### 4.1 考え方

環境汚染の測定局の適切な配置のあり方はモニタリングの目的に応じて検討されるべきであることは既に前章で述べた。そして、前章では、汚染対策の評価等のための地域全体の汚染状況の把握をモニタリングの目的とした場合において、必要最小な測定局の数と位置をどのように選定できるかを大気汚染データを対象に検討した。本章では水質汚濁データを対象に同様の検討を行う。松岡ら<sup>7)</sup>は統計的内挿法——データ間の関係を統計的に解析することにより内挿曲面を得る方法——を用い水質測定点の配置問題を検討している。本章では、内挿曲面を得る方法として前章と同様スプライン補間法を用いる。前章では一つの汚染物質（オキシダント）についての測定局の配置問題について検討したが、ここでは複数の項目を測定対象としている水質測定点の配置問題をとり扱う。また、削減可能測定点の選定についても前章と異なる考え方を採用する。これは、水質汚濁データでは同一測定点において測定されるべき対象が多いので、前章の方法では計算が膨大になり実用的でないからである。そのため、分布曲面を与える測定点の影響度から各測定点の相対的な重要性を評価する方法を提案し、これにもとづいて削減可能測定点及び代表測定点の選定を検討する。この方法の主要な考え方は次のようである。

- ① 対象とする濃度分布パターン：ここでは長期間の平均的濃度分布パターンの適切な把握を目的とした測定点配置問題を検討する。そのため、十分長い期間（ここでは1年間）の測定値の平均値による分布パターンを対象にとりあげる。
- ② 測定点の順位づけ：既存測定網の測定点には分布曲面形へ与える影響の強さに差がある。影響の強さによって測定点の順位づけをする。影響の強い測定点は重要性が高く、影響の弱い測定点は重要性が低いとする。

- ③ 削減可能性の検討等：重要性が低い測定点は削減が可能かどうかを検討する。削減の可能性の判断は分布曲面のへだたりの程度を判断基準として行う。重要性が高い測定点は代表的測定点——環境基準点又は自動測定機の設置点などとして採用しうる——の候補となる。
- ④ 複数項目についての結果の総合：水質汚濁の測定では同一の測定点で複数の水質項目が測定されるのが一般的であり、測定点配置が一つの水質項目にとっては適切であっても他の項目にとってはそうでない場合もあるので、総合的な評価を加えることにより、どの水質項目にもあてはまる結果を得るようにする。

## 4.2 測定点の分布曲面への影響度

配置された各測定点の濃度分布曲面形成へ与える影響を推定する。そして、これを判断基準として分布曲面への影響が最も小さくなるような測定点の抜き取り順位を決定する。

### (1) 影響度

各測定点のデータが濃度分布曲面に与えている影響の大きさを知るため、その測定点を抜き取ったときの分布曲面と元の曲面との差を尺度として採用する。全測定点のデータを用いて作成した分布曲面（基準曲面と呼ぶ）と任意の1測定点のデータを抜き取って作成した分布曲面（1点削減曲面と呼ぶ）の差

$$\frac{1}{M} \sum_{(X_{\alpha}, Y_{\beta}) \in A} |Z_{\alpha, \beta}^0 - Z_{\alpha, \beta}^1|$$

（ただし、 $M$ は湖水面に存在する格子点数、 $Z_{\alpha, \beta}^0, Z_{\alpha, \beta}^1$  はそれぞれ基準曲面、1点削減曲面の各メッシュ点でのデータ値）

を、基準面のそれ自身の平均値からのへだたりの程度を表わす値

$$\frac{1}{M'} \sum_{(X_{\alpha}, Y_{\beta}) \in A} |Z_{\alpha, \beta}^0 - \bar{Z}_{\alpha, \beta}^0|$$

（ただし、 $\bar{Z}_{\alpha, \beta}^0$  は基準曲面の各メッシュ点でのデータ値の平均値）  
で除し、百分率で表わしたものを影響度とし、次のように表わす。

$$G_0^1 = 100 \frac{\sum_{(X_\alpha, Y_\beta) \in A} |Z_{\alpha, \beta}^0 - Z_{\alpha, \beta}^1|}{\sum_{(X_\alpha, Y_\beta) \in A} |Z_{\alpha, \beta}^0 - \bar{Z}_{\alpha, \beta}^0|} \dots\dots\dots(6)$$

$G_0^1$ は基準曲面と1点削減曲面との差に基づく影響度を表わすが、一般的に、  
 j 個の測定点を抜きとって作成した分布曲面（j 点削減曲面と呼ぶ）と j' 点削減曲面との差に基づく影響度を次のように表わす。

$$G_j^{j'} = 100 \frac{\sum_{(X_\alpha, Y_\beta) \in A} |Z_{\alpha, \beta}^j - Z_{\alpha, \beta}^{j'}|}{\sum_{(X_\alpha, Y_\beta) \in A} |Z_{\alpha, \beta}^j - \bar{Z}_{\alpha, \beta}^j|}, j < j' \dots\dots\dots(7)$$

## (2) 抜き取り順位

測定点を順次抜き取ってゆく場合の分布曲面への影響の程度は抜き取りの順番によって異なる。影響度を最も小さくする順番を決定する。まず  $G_0^1$  を各測定点ごとに求め、最小の  $G_0^1$  の測定点を順位1の測定点とする。つぎにこの測定点を除き各測定点ごとに  $G_0^2$  を求め、最小の  $G_0^2$  の測定点を順位2の測定点とする。以下順次同様の手続を繰り返し、各測定点順位を決定する。この順位列（測定点順位列と称する。）のはじめに位置する測定点ほど濃度分布形成に関し重要性は低く、後方に位置する測定点ほどその重要性は高い。測定点順位列の決定を水質項目ごとに行う。

## (3) 実データへの適用

前述の琵琶湖のデータ（13項目、22測定点、53年度平均値（表1参照））に(1)、(2)の考え方を適用する。22測定点のそれぞれについて影響度（ $G_0^1$ ）を水質項目ごとに計算したものを一覧表とした（表2）。表から、測定点の影響度は水質項目によって若干異なるが、全体的には一定の傾向が見られる。たとえば、11番、12番、17番など湖面中央または東岸側測定点では影響度が大きい（10%以上）項目が多い。平均影響度（各測定点の水質項目ごとの影響度の幾何平均値）を図9に示した。平均影響度の大きい測定点は、湖面中央の18、11、7番測定点、東岸側の17、12番測定点（以上、平均影響度



表2 各測定点の分布曲面に与える影響度  $G_0^b$  (%)

Stn No.	Chl. a	SS	T-COD	P-COD	S-COD	T-N	Org-N	In-N	NO <sub>3</sub> -N	NH <sub>4</sub> -N	T-P	PO <sub>4</sub> -P	P-P	Geom. mean
1	2.18	2.34	2.99	4.99	0.104	6.54	6.52	1.83	7.39	2.76	1.80	4.01	2.36	2.57
2	3.92	1.99	2.79	4.90	1.73	3.81	3.98	0.226	3.32	2.36	0.169	18.4	1.84	2.20
3	4.96	4.14	4.70	8.38	1.80	3.90	5.05	5.99	3.09	7.61	2.49	9.77	1.63	4.28
4	1.31	2.63	1.92	2.87	1.82	1.35	1.43	0.292	1.84	1.66	3.33	1.14	3.48	1.67
5	0.793	1.92	0.821	0.309	2.04	2.35	2.32	0.639	1.69	0.0401	1.70	0.398	1.93	0.897
6	4.94	2.86	0.643	4.21	4.27	13.2	14.2	1.85	3.04	0.930	19.2	5.14	20.4	4.51
7	4.65	3.47	2.52	2.44	3.91	4.93	3.38	9.79	10.22	8.91	8.74	3.17	9.17	5.07
8	1.70	0.0263	2.35	0.738	4.57	2.59	3.77	6.22	3.71	7.77	2.88	7.46	3.95	2.31
9	2.15	3.54	3.55	2.21	1.03	4.28	3.46	6.11	1.14	9.40	2.30	0.302	2.47	2.43
10	1.22	4.05	1.39	1.17	1.03	1.11	0.854	1.60	0.775	2.21	4.85	4.16	5.74	1.82
11	6.19	0.172	5.27	8.66	3.85	6.35	4.83	10.7	10.56	9.75	11.2	17.0	10.4	5.98
12	17.5	9.05	11.2	12.8	7.59	9.26	3.96	32.7	19.3	40.9	13.2	18.3	12.4	13.5
13	0.732	8.07	0.0443	3.29	5.58	1.61	0.486	6.66	5.63	15.9	7.73	1.55	8.25	2.58
14	10.5	13.2	5.12	4.38	1.19	8.56	7.80	6.26	4.11	7.69	12.9	0.0321	14.1	4.47
15	0.429	6.94	1.56	8.06	7.45	2.07	1.83	1.16	0.245	2.15	0.497	5.06	1.02	1.77
16	1.44	14.8	4.32	6.24	4.91	9.98	10.8	2.75	6.60	0.456	1.24	4.00	0.906	3.49
17	11.4	19.9	9.75	14.7	1.48	8.22	4.87	21.2	19.5	20.5	14.4	0.826	15.6	9.12
18	4.37	13.6	2.38	6.76	15.2	5.43	4.51	6.47	6.29	5.49	6.97	2.80	7.29	5.93
19	0.343	2.34	1.46	0.0326	3.15	1.15	1.52	1.78	4.08	0.0564	0.0914	0.301	0.119	0.448
20	1.95	1.64	2.62	1.08	6.61	1.34	0.179	6.90	9.75	4.70	5.55	1.78	6.24	2.62
21	2.76	0.941	2.04	1.35	0.901	0.336	1.32	5.55	8.28	2.90	2.61	14.2	4.38	2.34
22	4.24	5.97	3.08	2.48	2.85	2.37	1.72	4.23	4.16	3.74	5.25	10.8	4.54	3.83

5以上)であり、平均影響度の小さいのは、西岸側の4, 5, 10, 15, 19番測定点(以上、平均影響度2以下)である。

測定点順位列を水質項目ごとに求めた(表3)。総合的な測定順位列を求めるため、13水質項目についての測定点順位の平均を算出し、その値が小さい方から順に並べた順位列(平均測定点順位列と称する。)を求めた。これを表の最右欄に示した。この順位列は13項目についての各測定点の重要

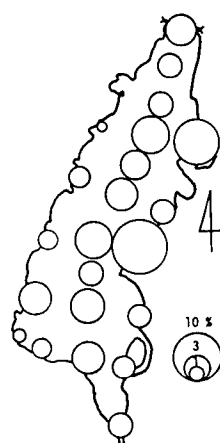


図9 各測定点の平均影響度(%)

性を表わす総合的な順位列であると言える。この場合、全水質項目についての測定点順位の単純平均を用いたが、場合によっては、特定の水質項目についてのみのそれを採用する方法または順位にウェイトをつけた加重平均を採用する方法も考えられる。モニタリングの目的、趣旨によって適当なものが選択されればよ

表3 測定点順位列

Order	Station Number													Ave. Order
	Chl. a	SS	T-COD	P-COD	S-COD	T-N	Org-N	In-N	NO <sub>3</sub> -N	NH <sub>4</sub> -N	T-P	PO <sub>4</sub> -P	P-P	
1	19	8	13	19	1	21	20	2	15	5	19	19	19	19
2	15	11	6	5	10	10	13	4	10	19	2	14	16	5
3	13	21	5	8	9	19	10	5	9	16	15	9	15	10
4	10	20	19	10	21	4	4	15	4	6	16	5	3	4
5	5	19	15	9	17	13	19	19	5	2	5	4	1	2
6	9	5	18	20	14	8	22	1	2	1	9	17	2	15
7	8	4	20	13	3	7	17	10	3	15	1	20	8	8
8	16	1	10	14	2	20	2	16	22	21	3	13	5	13
9	21	3	7	22	5	17	3	22	21	22	21	16	9	9
10	20	6	8	4	4	2	8	21	18	10	20	10	10	20
11	1	10	4	6	16	3	9	6	19	4	10	7	21	3
12	3	2	11	15	19	6	15	8	13	3	8	1	20	21
13	6	7	1	16	12	5	18	3	14	8	13	6	13	16
14	4	12	3	1	11	15	16	13	8	18	18	8	18	1
15	2	9	2	3	13	22	14	9	6	17	7	2	7	6
16	11	16	22	2	6	16	6	17	1	20	6	3	6	17
17	17	18	17	17	7	18	5	20	16	9	4	15	4	22
18	22	17	16	21	22	11	7	14	17	12	17	21	17	7
19	18	22	21	7	18	12	12	12	20	7	22	22	22	18
20	12	14	12	12	20	14	21	7	12	11	14	18	14	14
21	7	13	9	11	8	1	11	11	7	13	11	11	11	11
22	14	15	14	18	15	9	1	18	11	14	12	12	12	12

い。表から、平均順位列のはじめに位置する測定点は、西岸側の 5、10、19 番測定点などであり、後方に位置するのは湖面中央の 7、11、14、18、22 番測定点、東岸側の 12、17 番測定点などである。

影響度 ( $G_0^1$ ) の大きい測定点は一般に測定点順位列の後方に位置し、 $G_0^1$  と測定点順位との間には概ね相関がある。

### 4.3 測定点削減による分布曲面の変化

水質項目ごとに決定された測定点順位に従って、測定点を抜き取り、残りの測定点のデータにより分布曲面を作成し、基準曲面とのへだたりを調べる。図 10 は、Chl.a、T-COD、T-N、T-P、SS について N 点削減曲面が、N が大きくなるに従い基準曲面からどのようにへだたってゆくかを見たものである。視覚による分布曲面のへだたりの程度を定量的な値と関連づけるため各分布曲面に  $G_0^N$  の値を記した。図から、基準曲面からの各分布曲面のへだたりは  $G_0^N < 10(\%)$  ではほとんど見られないが、 $G_0^N > 30(\%)$  ではかなり顕著であることがわかる。図 11

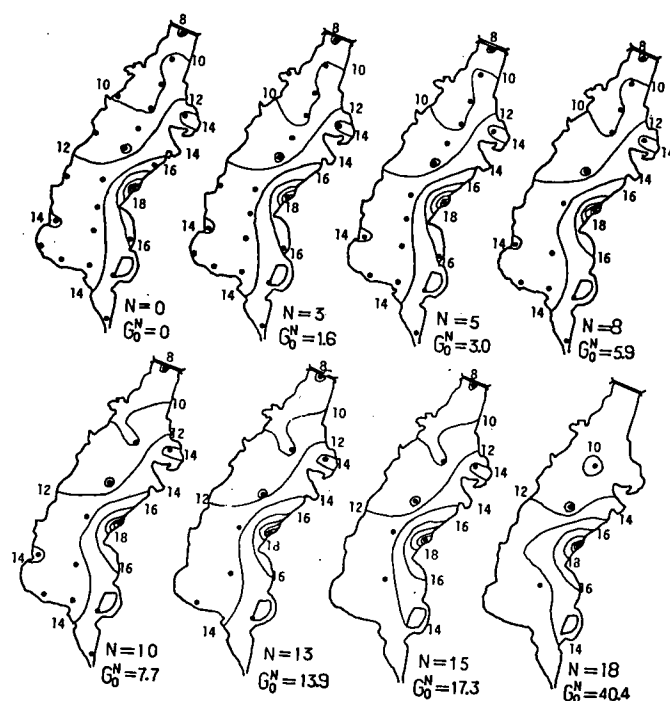


図 10(a) 測定点削減による分布曲面の変化 (Chl.a (mg/l))

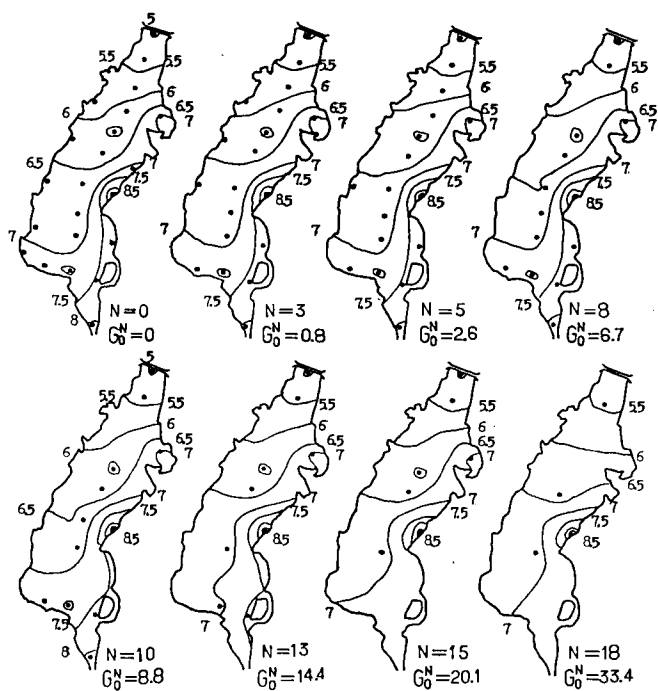


図 10(b) 測定点削減による分布曲面の変化 (T-COD (mg/l))

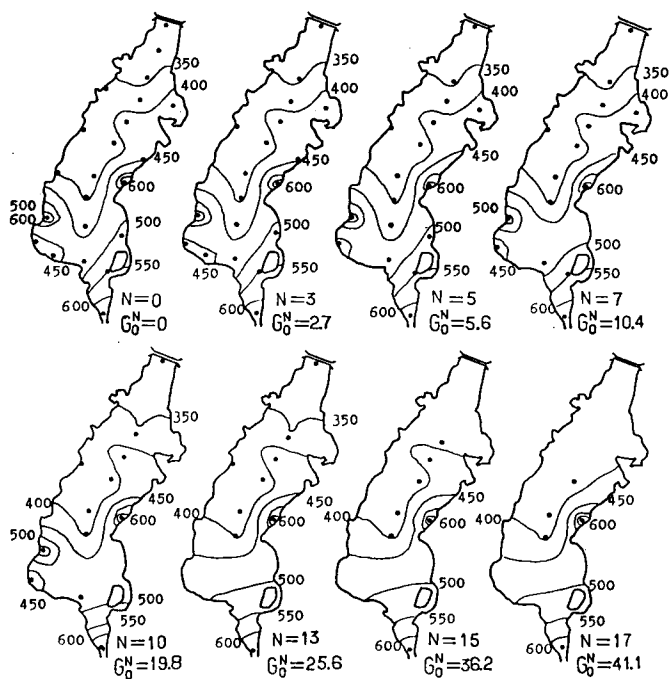


図 10(c) 測定点削減による分布曲面の変化 (T-N ( $\mu\text{g/l}$ ))

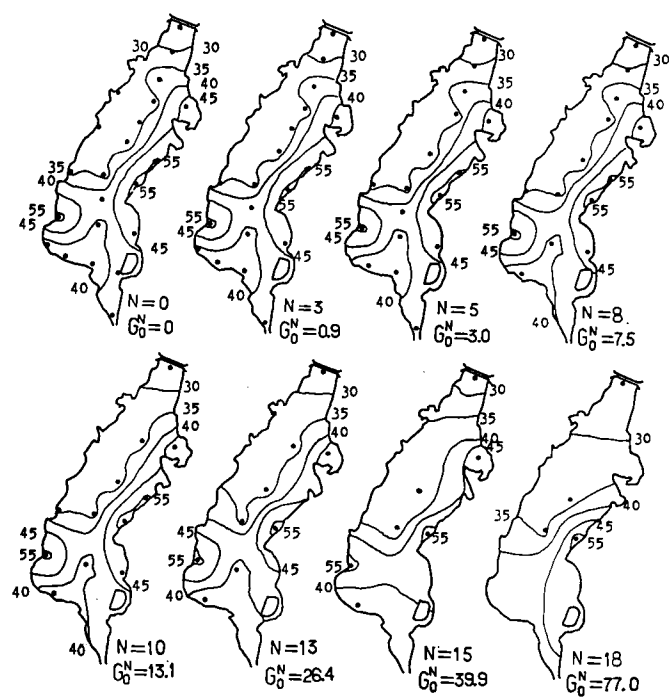


図 10(d) 測定点削減による分布曲面の変化 ( T-P (  $\mu\text{g}/\ell$  ) )

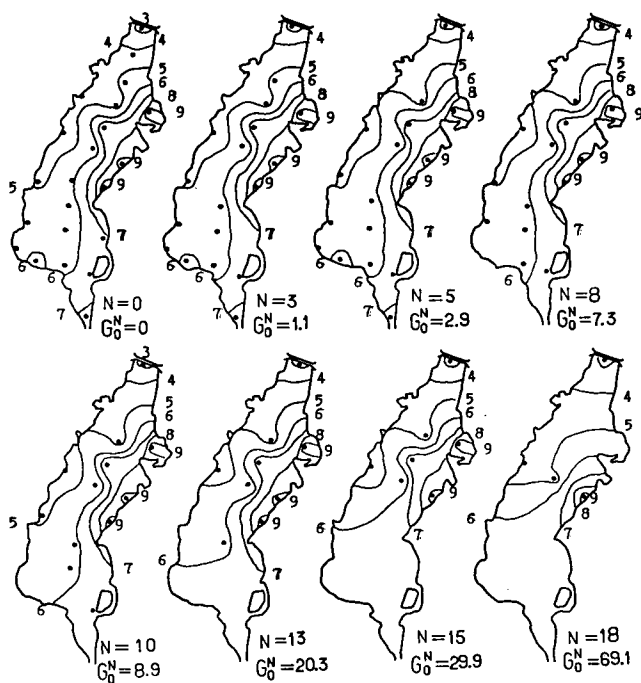


図 10(e) 測定点削減による分布曲面の変化 ( SS (  $\text{mg}/\ell$  ) )

に  $N$  による  $G_0^N$  の変化を Chl. a、T-COD、T-N、T-P、SS について示した。抜き取り測定点数 ( $N$ ) が 7 程度までは  $G_0^N$  の値は小さく (10% 以下)、その上昇カーブはゆるやかである。しかし、15 を起えると  $G_0^N$  の値は 20~40% 以上となり、その上昇傾向はきわだって大きくなる。図 1 2 に  $G_{N-1}^N$  について示した。 $N \leq 6$  では項目にかかわらず  $G_{N-1}^N$  は小さい値 (5% 以下) である。 $N > 10$  では、 $N$  によりまた項目により、その値は 4%~50% の範囲で大きくふれるが、一般的には増加傾向が顕著になる。表 4 に、 $G_0^N < 10\%$  を満足する  $N$  の最大値を項目ごとに示した。同時に、 $G_{N-1}^N < 4 \sim 7\%$  を満足するそれを並記した。条件  $G_0^N < 10\%$  と条件  $G_{N-1}^N < 4 \sim 7\%$  は概ね同一の結果を与えることがわかる。

なお、収束条件  $\eta_m < 10^{-4}$ 、

$\eta_s < 10^{-5}$  にかかわらず、繰り返

し回数 200 回で計算を打ち切ったが、これによる影響は、抜き取り測定点数が 17 以上 (残り測定点数が 5 以下) になった場合にのみ生じており、本論文における考察にはほとんど影響を及ぼすものでないことが示された。

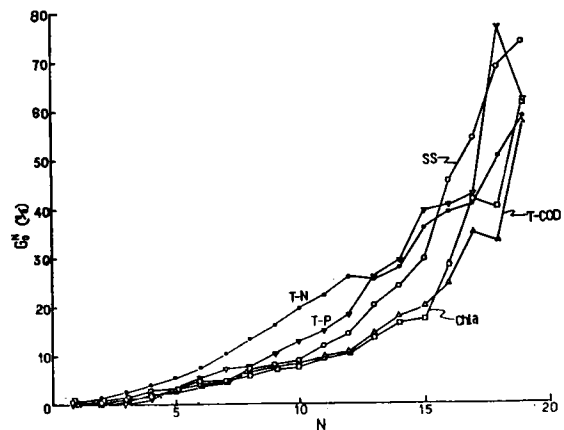


図 1 1  $G_0^N$  と  $N$  との関係

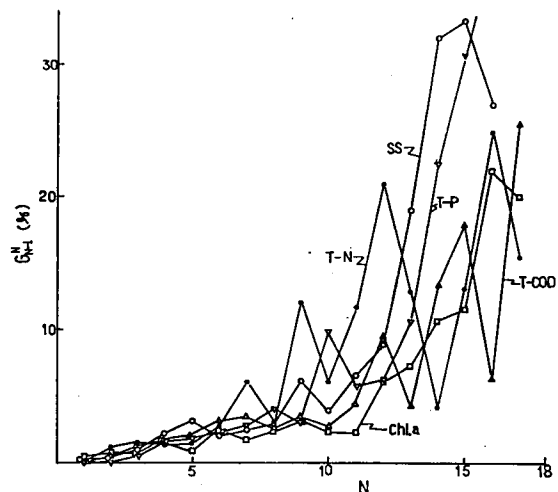


図 1 2  $G_{N-1}^N$  と  $N$  との関係

表4  $G_0^N < 10\%$ 、 $G_{N-1}^N < 4 \sim 7\%$ を  
満足するNの最大値

	$G_0^N < 10\%$	$G_{N-1}^N < X$			
		X=4	X=5	X=6	X=7
Chl. a	11	11	11	11	12
SS	10	5	7	7	7
T-COD	10	10	11	11	11
P-COD	8	7	10	10	10
S-COD	11	10	10	10	10
T-N	6	8	6	6	8
Org-N	7	6	6	6	8
In-N	8	7	8	10	11
NO <sub>3</sub> -N	7	6	6	8	8
NH <sub>4</sub> -N	8	8	8	8	8
T-P	8	5	5	6	6
PO <sub>4</sub> -P	9	8	8	8	8
P-P	8	6	6	6	9

#### 4.4 削減可能測定点の選定及び代表測定点の選定

削減可能測定点の選定の具体的方法は次のようである。まず、全測定点にもとづく分布曲面を基準曲面とし、対象とする水質項目（複数）についての測定点の平均測定点順位列を決定する。いずれの水質項目についてもこの順位に従って順次測定点を抜き取る。水質項目ごとに分布曲面の基準曲面からのへだたり  $G_0^N$  を計算する。 $G_0^N$  がいずれの項目においても一定未満——この条件は前述のように  $G_0^N < 10(\%)$  が適当と考えられるが、測定の精度、目的により、事例に応じて適切に設定されればよく、また  $G_{N-1}^N$  の条件を加えることも考えられる。——であるNの最大値を求めれば、それが削減可能測定点数である。

実データ（前述のデータ）を適用し、削減可能測定点の選定を行う。対象水質項目（13項目）の平均測定順位列は表3に示されている。この順位列にもとづいて測定点を抜き取ってゆく場合の水質項目ごとの  $G_0^N$  の計算結果を表5に示す。表から、いずれの水質項目においても  $G_0^N < 10(\%)$  となるのは  $N=1 \sim 3$  の場合であることがわかる。したがって、削減可能測定点数は3、すなわち、19、5、10番測定点が削減可能測定点である。

また、対象水質項目として主要な5項目（Chl.a, T-COD, T-N, T-P, SS）のみをとりあげた場合は次のようである。5項目についての平均測定点順

表5 平均測定点順位列(13項目)にもとづく $G_0^N$

N	Sta. No.	G <sub>0</sub> <sup>N</sup> (%)													
		Chl. a	SS	T-COD	P-COD	S-COD	T-N	Org-N	In-N	NO <sub>3</sub> -N	NH <sub>4</sub> -N	T-P	PO <sub>4</sub> -P	P-P	MAX
1	19	0.3	2.3	1.5	0.0	3.2	1.2	1.5	1.8	4.1	0.1	0.1	0.0	0.1	4.1
2	5	1.1	4.2	2.2	0.3	5.1	3.5	3.8	2.4	5.6	0.1	1.8	0.4	2.0	5.6
3	10	2.2	8.2	3.5	1.5	5.9	4.5	4.6	3.9	6.2	2.3	6.5	4.5	7.7	8.2
4	4	5.2	8.0	4.5	5.5	5.4	9.5	9.6	5.1	5.9	5.0	14.1	5.0	15.7	15.7
5	2	7.9	10.0	7.5	10.9	7.0	11.3	11.6	5.4	8.8	6.4	14.3	22.7	18.1	22.7
6	15	8.0	21.9	8.2	23.0	17.4	15.1	15.1	5.8	9.6	10.0	14.5	26.0	18.1	26.0
7	8	11.4	22.7	12.5	25.9	20.0	19.5	21.2	14.5	15.7	16.2	16.8	24.6	19.2	25.9
8	13	12.2	30.2	12.5	29.2	24.7	20.8	21.4	21.7	20.4	32.8	24.4	25.8	27.2	32.8
9	9	11.9	33.4	11.8	28.1	25.0	23.8	23.3	31.4	22.3	47.7	24.7	26.0	27.5	47.7
10	20	14.3	33.2	14.9	27.9	34.2	25.8	23.9	37.4	26.8	53.7	31.1	28.3	35.0	53.7
11	3	16.9	34.9	11.7	21.5	38.4	28.7	27.1	34.5	25.0	54.6	39.7	27.0	44.6	54.6
12	21	16.4	36.0	17.2	22.8	46.4	30.9	30.3	32.7	24.8	53.4	39.2	42.3	42.9	53.4
13	16	17.6	41.4	23.5	21.3	59.3	31.5	34.0	38.5	24.5	57.4	39.3	44.1	43.7	59.3
14	1	17.5	42.1	32.1	35.1	61.0	38.8	40.3	45.2	38.2	55.6	40.7	43.5	45.0	61.0
15	6	17.3	41.4	34.3	33.2	69.9	53.4	58.5	47.0	38.7	60.5	42.9	49.4	45.1	69.9
16	17	33.3	82.5	33.1	49.0	76.6	66.0	66.6	56.7	59.9	63.2	75.2	53.1	78.1	82.5
17	22	29.9	74.8	51.7	47.1	116.3	66.1	68.5	53.3	60.4	63.0	66.6	40.8	71.3	116.3
18	7	55.6	106.3	49.7	55.0	115.6	60.5	82.3	151.4	116.2	189.3	64.0	40.2	69.9	189.3
19	18	80.7	113.1	58.6	82.9	79.5	60.8	83.8	196.2	128.1	250.5	62.3	81.4	66.6	250.5
20	14	197.4	148.9	121.4	143.8	101.7	89.2	90.0	312.4	181.0	410.6	117.5	138.1	116.0	410.6

表6 平均測定点順位列(5項目)にもとづく $G_0^N$

位列を求め(表6の第2欄  
に示した)、これにもとづ  
いて各項目についての $G_0^N$   
を計算した(表6)。同様  
にして削減可能測定点数と  
して4が導き出される。す  
なわち、19、5、10、  
8番測定点が削減可能測定  
点である。

対象水質項目の数によっ  
て上述のように結果が異な  
るが、対象項目数が13と

N	Sta. No.	$G_0^N$ (%)					
		Chl. a	T-COD	T-N	T-P	SS	MAX
1	19	0.3	1.5	1.2	0.1	2.3	2.3
2	5	1.1	2.2	3.5	1.8	4.2	4.2
3	10	2.2	3.5	4.5	6.5	8.2	8.2
4	8	3.9	5.8	7.0	9.2	8.2	9.2
5	20	6.2	9.0	8.7	15.7	8.3	15.7
6	21	5.7	14.7	10.5	15.1	8.6	15.1
7	13	6.5	14.8	11.8	22.7	16.6	22.7
8	15	6.5	14.9	16.1	22.7	28.3	28.3
9	4	9.3	16.0	20.7	29.9	28.1	29.9
10	6	11.2	18.2	23.7	35.1	27.8	35.1
11	2	16.8	23.4	29.4	35.9	30.9	35.9
12	3	15.3	18.0	26.1	35.3	29.0	35.3
13	1	17.5	31.5	53.2	35.4	33.2	53.2
14	16	18.6	37.7	57.4	35.6	39.6	57.4
15	7	23.0	35.8	59.0	44.5	42.5	59.0
16	11	36.4	36.8	58.5	63.0	46.1	63.0
17	9	104.9	65.5	51.8	97.9	107.6	107.6
18	18	141.8	78.4	61.0	115.9	139.5	141.8
19	17	131.6	67.9	60.2	131.5	150.3	150.3
20	22	160.6	101.1	71.6	130.0	158.0	160.6

5と大きく異なる場合でも、結果は8番測定点を削減対象とするか否かのための相違にすぎない。

また、削減可能測定点の検討結果は、5.1(3)の平均影響度についての議論から得られた結果——平均影響度の小さい測定点は4、5、10、15、19番測定



点である——と概ね同様であることが知られる。

全測定点（22個）のデータを用いて作成した分布曲面（ $N=0$ ）と上記の3測定点のデータを用いないで作成した分布曲面（ $N=3$ ）とを比較した。図13は代表的な水質項目について両者を対比したものである。両者は視覚的にはほとんど同一の分布曲面に見える。したがって、削減可能測定点は19、5、10番測定点であるという上述の検討結果は妥当である。

基準曲面を変更した（上記3測定点を取り除いた19測定点による分布曲面を新たに基準曲面とした）場合の結果を検討する。新たな基準曲面（3点削減曲面）ともう1測定点削減した場合の分布曲面（4点削減曲面）のへだたり $G_3^4$ は8.1% < 10%（表7より）であるので前述の3測定点に加えてもう1測定点（4番測定点）が削減可能測定点となる。しかし、 $G_{N-1}^N$ は $N=5$ 以上いかなる $N$ につい

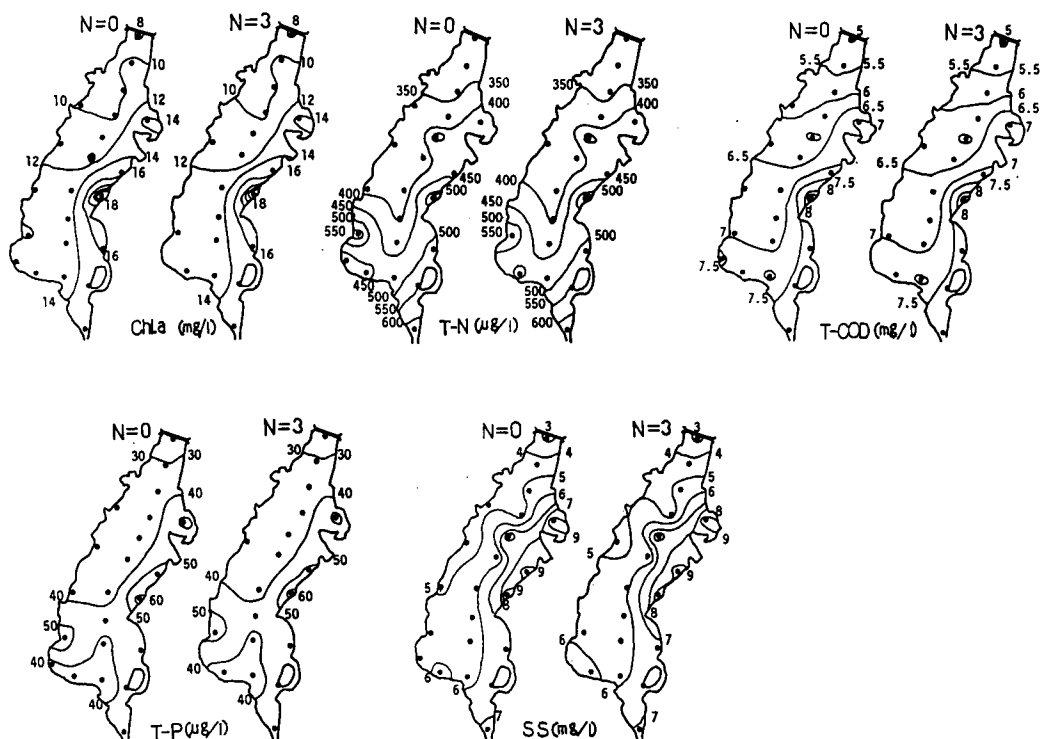


図13 基準曲面と削減可能測定点抜き取り後の分布曲面との比較

表7 平均測定点順位列(13項目)にもとづく $G_{N-1}^N$

N	Sta. No.	G <sub>N-1</sub> <sup>N</sup> ( % )													
		Chl. a	SS	T-COD	P-COD	S-COD	T-N	Org-N	In-N	NO <sub>3</sub> -N	NH <sub>4</sub> -N	T-P	PO <sub>4</sub> -P	P-P	MAX
1	19	0.3	2.3	1.5	0.0	3.1	1.2	1.5	1.8	4.1	0.1	0.1	0.0	0.1	4.1
2	5	0.8	2.0	0.8	0.3	2.1	2.4	2.4	0.7	1.7	0.0	1.7	0.4	1.9	2.4
3	10	1.2	4.1	1.4	1.1	1.0	1.1	0.8	1.6	0.8	2.2	4.8	4.2	5.6	5.6
4	4	3.1	2.5	2.4	4.6	0.9	4.9	5.0	1.2	1.3	2.8	7.6	1.5	8.1	8.1
5	2	3.9	2.3	3.1	5.6	1.8	3.5	3.7	0.3	3.3	2.4	0.3	19.2	2.3	19.2
6	15	0.5	12.8	2.2	13.0	12.1	3.7	3.4	1.4	1.2	3.8	0.3	8.1	0.4	13.0
7	8	3.8	0.9	4.6	3.0	6.5	5.0	6.7	9.0	6.2	10.9	3.8	4.8	4.4	10.9
8	13	0.9	10.0	0.1	3.6	6.8	1.7	0.4	6.8	5.4	16.5	8.1	2.0	8.5	16.5
9	9	2.6	7.4	4.5	2.4	1.2	6.1	4.7	8.7	2.4	12.2	3.0	0.9	3.1	12.2
10	20	2.6	1.6	3.7	1.0	11.1	2.1	0.7	7.4	11.4	4.9	7.4	3.1	8.2	11.4
11	3	13.9	10.9	7.7	15.0	8.5	12.7	15.4	11.0	3.7	14.6	14.3	15.8	14.0	15.8
12	21	3.4	2.5	6.2	2.5	11.7	2.1	2.9	5.2	9.4	1.4	0.6	27.9	3.0	27.9
13	16	2.3	33.6	8.6	7.9	21.3	15.4	16.4	6.1	9.7	3.4	0.1	6.2	0.8	33.6
14	1	5.8	10.7	15.1	21.3	3.5	20.9	19.2	9.5	28.9	3.2	1.5	3.5	1.4	28.9
15	6	9.5	3.1	4.1	11.1	25.3	39.7	41.3	7.6	6.8	8.4	43.7	15.0	46.3	46.3
16	17	22.7	63.5	20.4	31.0	14.9	27.5	19.9	31.6	36.9	26.2	41.7	5.9	46.0	63.5
17	22	12.4	24.0	32.8	10.2	92.1	15.3	14.9	6.4	1.0	10.6	22.4	25.0	21.7	92.1
18	7	48.6	70.8	31.6	31.8	5.5	12.2	24.5	138.3	135.9	137.4	52.0	27.6	54.9	138.3
19	18	31.8	44.2	19.6	57.4	103.7	27.1	20.1	51.2	66.1	45.3	21.8	55.5	17.7	103.7
20	14	110.7	111.7	111.6	107.2	111.0	107.7	103.6	109.7	103.0	112.7	112.7	95.4	112.9	112.9

でも10%以上となることはないので、基準曲面変更による結果の差は4番測定点が削減可能となるか否かの相違にすぎない。なお、分布曲面のへだたりの程度の判断基準として $G_0^N$ に加えて $G_{N-1}^N$ をも用いれば、このような相違の生じることが抑制される(本例の場合では $G_{N-1}^N < 6 \sim 7$ が適当である)。

つぎに、少数の代表的な測定点の選定方法を考える。水域を代表する測定点(1～数点)の選定は、環境基準点をどこに選ぶか、自動測定機をどこから設置し始めればよいかなどを検討する際に役に立つ。この場合においては、平均測定点順位列を求め、これの後方に位置するいくつかの測定点の中から、そのとるデータ値や地理的条件等を勘案して、適切なものを選び出せばよい。前述の例では12、11、14、18、7番測定点が該当する。ちなみに、現在、11番測定点が環境基準点として採用されており、OECDの調査では14番測定点が基準測定点として採用されている<sup>8)</sup>。

## 第5節 結 語

本章では、スプライン補間法を湖の水質濃度分布パターンの推定に適用するための条件及び技術を検討したのち、これを実際に琵琶湖南湖の実データに適用し、その実用性を実証した。ついで、スプライン補間法により作成される分布曲面に

与える測定点の影響から各測定点の相対的な重要性を評価することにより削減可能測定点及び代表測定点の選定のし方について検討した。得られた主要な成果は以下のようであった。

(1) 補間計算を精度よくかつ迅速にできるようにするための条件及び技術を検討し、次のような結果を得た。

- ① 逐次近似計算では、逐次誤差最大値( $\epsilon_m$ )、逐次誤差平均値( $\epsilon_s$ )がそれぞれ  $10^{-4}$ 、 $10^{-5}$  未満になれば収束したとみなすことで実用上十分である。
- ② 水質汚濁の測定のようにデータ点が少ない場合はスプライン補間の収束性は悪いが、この向上を図るためSOR(Successive Over-relaxation)法を導入することは効果がある。SOR法の加速パラメータ( $\omega$ )は、測定点が比較的多い場合は1.7～1.9程度が適当であるが、測定点が極端に少ない場合は2.0に近いほどよい。通常の場合は $\omega = 1.9$ 付近を用いれば十分である。
- ③ スプライン・パラメータ( $sh^2$ )は、ほとんど全ての水質項目について、計算結果の誤差に顕著な差をもたらさないが、いくつかの理由—— $sh^2$ が大きくなるほど計算の収束性はよくなる。 $sh^2$ が小さいほど作成される補間分布曲面は滑らかであると言われる。大気汚染データの場合のこの値は1～10が適当である。——から水質汚濁データについても1～10を用いれば十分である。

(2) 水質濃度分布パターン及びその季節的変化の視覚的把握にスプライン補間法の適用が有効であることを、琵琶湖南湖の実データにより示した。

(3) 削減可能測定点及び代表測定点の選定方法について検討し、次のような結果を得た。

- ① スプライン法により作成される分布曲面へ与える測定点の影響度を、基準曲面とN点削減曲面との差( $G_0^N$ )によって表現し、これをもとに測定点順位列を決定し、各測定点の相対的な重要性を評価する考え方を提案した。また、琵琶湖南湖の例で、相対的重要性が高い測定点を具体的に示した。
- ② 測定点削減による分布曲面の変化が視覚的にはほとんど認められない影響度  $G_0^N$  の条件は、 $G_0^N < 10(\%)$ であった。

- ③ 琵琶湖南湖に設置された 22 個の測定点の例において削減可能な測定点と代表的測定点の選定を行った。全水質項目に共通して影響度の小さい 3 測定点（19、5、10 番測定点）が削減可能測定点として選定された。対象水質項目の数又は基準曲面を変更すると結果は変わってくるが、その変化の程度はわずかであった。平均順位列の後方に位置するいくつかの測定点（11、14、18、7 番測定点等）が代表的測定点の候補として選定された。

## 結 論

本論文では、環境データの特徴に起因して生じるデータ処理、利用上の問題点を指摘し、そのうち主要ないくつかの課題を検討対象としてとりあげた。まず、第Ⅰ編（第1章～第4章）において環境データの組織化の意義とその具体的方策について論じ、第Ⅱ編（第5章～第8章）において環境データを環境問題へ適用するために必要な技術手法（データチェック方法、データの2次元化方法、測定点の最適配置方法を取りあげた）を論じた。本研究の主要な成果を各章ごとに要約する。

第1章は、大気環境データを組織化するための方策について論じたものである。まず、データの観測、移送、チェック、編集、入力、蓄積、管理の各段階ごとに課題を指摘し、とくに、わが国における大気環境データシステムの現状から主たる課題を抽出した。磁気テープへのデータの記録方式の標準化については具体案を提案した。

第2章は、水質環境データと廃棄物関係データを組織化するための方策について論じたものである。第2節ではわが国における水質環境データシステムの現状と課題を述べるとともに、磁気テープへのデータの記録方式の標準化については具体案を提案した。第3節では廃棄物関係データの制度的な把握の実態とコンピュータによる管理の実例をレビューし、将来の廃棄物データシステムの有すべき機能について考察した。

第3章は、国内外における環境情報システムの開発、運営状況をレビューしたものである。外国については、米国、カナダ、西独、スウェーデンの例を取りあげた。とくに、米国における大気環境データを対象とする SAROAD、水質環境データを対象とする STORET などの全国的規模のデータベースの内容は今後のわが国の環境データベースの開発に関し、有益な示唆を与えるものとして特筆した。

第4章は、組織化されたデータにもとづきデータベースを作成することにより、データの処理、利用がどのように便利になるかを論じたものである。まず、環境データベースを作成するにあたっての基本的な考え方を論じ、ついで、第1～2

章で検討したデータの組織化の内容にもとづき、データファイルを作成し、かつ必要なソフトウェアを開発することによりデータベースを作成した。このデータベースによりデータの処理や解析が便利に効率的に行えるようになることを実例（端末機からの指示によりデータが所定の内容の処理をされ所定の様式で出力される例や、汚染度が2次元分布パターンとして表現される例など）で示した。

第5章は、データの自動チェック・システムに関するものであり、まず、全国の自治体における大気汚染データのチェックの方法を概観し、現状の人手による方法の問題点を指摘した。これに対する対応策としてコンピュータを用いたデータチェック法をとりあげた。データチェック法の比較、評価のための方法を提案し、これにもとづき、コンピュータによる各種の自動チェック法の比較、検討を行った。その結果、時系列値チェック法と上下限值チェック法の組合せが比較的効率のよい異常値検出法であることを得、その実用性を示した。リアルタイムにデータをチェックする方法について、ADF(Adaptive Digital Filter)法をとりあげて検討し、概ね良好な結果を得、実用化の可能性を示した。

第6章は、離散的な限られた地点で測定された大気汚染データを2次元の分布パターンとして表現するためのデータ補間法を論じたものである。データ補間法としてスプライン法を採用した。スプライン法が大気汚染データの補間に良好な結果を与えるかどうかを検討するため、スプライン法による補間分布曲面が真の分布曲面の近似分布曲面と見なせるかどうかを判断する方法を提案した。測定局数の増加による補間分布曲面の近似分布曲面への接近の程度を表わす量  $D_k^*/R_k^*$  と、補間分布曲面の一定曲面形状への収束の挙動を表わす量  $Y_k^*$  を判断基準として導入した。この方法を実際の大気汚染データに適用することにより、測定局数が一定以上に増大すると  $D_k^*/R_k^*$  は小さな値をとりつづけること、及び  $Y_k^*$  が  $k$  の増大に伴って一定値に収束することを示し、スプライン法の大気汚染データの補間への適用性が良好であることを実証した。また、この手続きにより、近似分布を得るに必要な最小の測定局数が代表的な分布パターンごとに求められることを述べた。

第7章は、既存の大気汚染測定網から取除くことが可能である測定局を選定す

る方法を論じたものである。この方法は、前章で論じた必要な測定局の数を求める方法に基づき、取除き可能な局を準最適に選定するものである。すなわち、できるだけ少ない測定局数で、できるだけ真の濃度分布に近い濃度分布結果が得られるようにする観点からの測定局の最適配置を検討した。削減すべき測定局の選定の方法として、濃度分布曲面の曲面形への影響が小さいと見られる局を抽出（分布曲面の曲率の大きさを判断基準とする）し、それを確認（補間分布曲面の元のそれからのへだたりを調べる）する方法を提案した。この方法を実データに適用し、削減が可能な測定局を具体的に選定した。全測定局のデータを用いて作成した補間分布曲面と削減可能と指定された測定局のデータを用いないで作成したそれとを比較し、両者の分布曲面の形状に大きな差がないことを示し、このことにより、提案した削減可能局選定方法が妥当であることを実証した。

第 8 章は、水質汚濁濃度分布パターンの推定にスプライン法によるデータ補間法の適用を検討するものである。検討の内容は前二章と概ね同様であるが、水質汚濁データの特徴を考慮し、データ補間計算をする上でのいくつかの条件（収束条件、収束性の向上措置、スプライン・パラメータ）の最適値域を設定した。スプライン補間法を琵琶湖湖南湖の水質測定データに適用し、2 次元濃度分布パターン及びその季節的変化が良好に表現できることを示した。測定点の合理的配置の検討は、測定点が多くの水質項目を対象としている点に留意し、次のような考え方に基づいて行った。スプライン法により作成される分布曲面へ与える測定点の影響度を、基準曲面と N 点削減曲面との差 ( $G_0^N$ ) によって表現し、これをもとに測定点順位列を決定し、各測定点の相対的な重要性を評価する考え方を提案した。琵琶湖湖南湖に設置された 22 個の測定点の例において、この考え方にもとづいて、削減可能測定点として 3 測定点を、代表的測定点の候補としていくつかの測定点を選定した。

以上が本論文の主たる成果である。今後、環境データのコンピュータによる総合的な管理が進むと考えられ、データの組織化の努力が一層要請されることとなろう。本論文では、この問題に関し、環境データのうち代表的な大気環境データ、

水質環境データ、廃棄物データを対象として課題と今後の方策について検討した。また、このようなデータを環境問題に適用するに必要な2～3の技術手法を検討した。環境データの範囲は広くかつ用いるべき技術手法也多岐にわたるので今後に残された課題も多い。これらについては、本研究の成果を基礎として、引続き検討してゆきたいと考えている。

終りに臨み、本論文の作成にあたり、終始懇篤な御指導、御鞭撻を賜った京都大学教授 宗宮 功博士、同平岡正勝博士に対し深甚の謝意を表します。また、本研究の遂行にあたって、御指導、御援助を頂いた国立公害研究所水質土壤環境部長 合田 健博士を始め、同研究所総合解析部長 内藤正明博士、環境情報部長 廣崎昭太博士、情報システム室長 大井 紘博士、徳島大学工学部教授 添田 喬博士、京都大学工学部助教授 津野 洋博士、同助手 藤井滋穂氏に心から感謝の意を表します。また、本研究に御協力いただいた松本幸雄、宇都宮陽二郎、大島高志、新藤純子の諸氏には心よりお礼申し上げます。なお、本論文がもとになっている情報科学や環境行政の知識は、著者が席を置いた環境庁、厚生省、国立公害研究所の各部局の関係者に負うところ大であり、暖かい御支援をいただいた厚生省水道環境部長 山村勝美博士、国立公害研究所長 近藤次郎博士を始め、その他関係各位に感謝いたします。

最後に、本論文で使用したデータは関係自治体及び研究機関から提供していただいたものであること、これの計算、解析は国立公害研究所及び京都大学大型電算機センターのコンピュータを用いて行ったことを付言します。



## 使 用 記 号

### 第 5 章

- $N/N_0$  データチェック手法による異常値の発見の割合を表わす量で、発見率と呼ぶ。総数  $N_0$  個の異常値のうち  $N$  個が発見される場合を表わす。
- $N/M$  データチェック手法による異常値の適中の割合を表わす量で、適中率と呼ぶ。検出した  $M$  個の疑問値のうち  $N$  個が真の異常値である場合を表わす。

### 第 6 章及び第 7 章

- $A$  対象とする領域（領域  $A'$  の四囲から境界条件を与えるためのメッシュの 2 列分を除いた長方形領域）
- $A'$  対象とする領域  $A$  を含む適当に広い領域
- $b$   $Y_k^s$  の  $k$  による変動をみる場合の基準となる分布曲面の測定局数
- $c_m$   $m_k$  が十分小さい値と判断する場合の判定基準
- $c_s$   $m_k + s_k$  が十分小さい値と判断する場合の判定基準
- $D_k^s$   $P_k^s(x, y)$  と  $P_k^{s*}(x, y)$  とのへだたりの尺度、具体的には、  

$$\frac{1}{A} \iint_A |P_k^s(x, y) - P_k^{s*}(x, y)| dx dy$$
- $D_k^s / R_k^s$  補間曲面が近似曲面に近いかどうかを判断するための関数
- $[F]$  関数  $F(x, y)$  の  $A$  上での平均値、具体的には、  

$$\frac{1}{A} \iint_A F(x, y) dx dy$$
- $k'$   $D_k^s / R_k^s$  が小である  $k$  の区間のうち最小値（整数）
- $k''$   $D_k^s / R_k^s$  が小である  $k$  の区間のうち  $k'$  より十分大である  $k$ （整数）
- $[k', k'']$  整数  $k'$  から整数  $k''$  までの整数  $k$  の区間
- $k'_i$  時刻  $t_i$  における分布曲面についての  $k'$

$m$	$i \in n$ のうち $[k', k']$ の中で $D_k^{s(i)} / R_k^{s(i)}$ の値が $m_k + s_k$ より小さい $i$ の集合及び集合の要素数
$m_k$	各 $k$ での $D_k^{s(i)} / R_k^{s(i)}$ の平均値、具体的には、 $\frac{1}{n} \sum_{s(i) \in n} D_k^{s(i)} / R_k^{s(i)}$
$N$	A 内に現存する測定局総数
$n$	N 個の測定局から N 個の測定局を選び出す全ての順列から、ランダムに選んだ順列の集合及びその要素数
$P(x, y)$	真の濃度分布曲面
$P_s(x, y)$	$P(x, y)$ の近似分布曲面
$P_k^s(x, y)$	$s$ なる順列のうち $k$ 番目までの測定局データにより補間した曲面
$P_k^{s*}(x, y)$	$P_k^s(x, y)$ の比較曲面、具体的には、 $P_{k+r}^s(x, y)$
$R$	$P(x, y)$ のその平均値 $[P]$ からのへだたりの尺度、具体的には、 $\frac{1}{A} \iint_A  P(x, y) - [P]  dx dy$
$R_k^s$	$P_k^s(x, y)$ のその平均値 $[P_k^{s*}]$ からのへだたりの尺度、具体的には、 $\frac{1}{A} \iint_A  P_k^{s*}(x, y) - [P_k^{s*}]  dx dy$
$r$	$P_k^s(x, y)$ の比較曲面として、分布曲面を、 $k$ を増加させて作成したときの $k$ の増加分
$s$	N 個の測定局から N 個の測定局を選び出す順列のひとつ
$s(i)$	N 個の測定局から N 個の測定局を選び出す順列のうち $i$ 番目の順列
$s_k$	各 $k$ での $D_k^{s(i)} / R_k^{s(i)}$ の標準偏差
$t_i$	十分長い期間の中の測定がされた全ての時刻のうち汚染濃度が高い時刻として選び出された時刻
$Y_k^s$	$k$ に伴う補間分布曲面の一定曲面形状への収束の挙動を表わす関数

Y 元の分布曲面  $P_0(x, y)$  と測定局を 1 局削減した場合の分布曲面  $P_1(x, y)$  のへだたりの尺度、具体的には

$$\iint_A |P_0(x, y) - P_1(x, y)| dx dy / \iint_A |P_0(x, y) - \{P_0\}| dx dy$$

$\sigma$  スプライン・パラメータ、弾性板の張力による伸びのエネルギーと歪エネルギーの係数の比

## 第 8 章

A 湖水面の領域

A' 計算対象領域

$G_0^N$  基準分布曲面と N 個の測定点を抜きとって作成した分布曲面とのへだたりの尺度、具体的には、

$$100 \frac{\sum_{(X_\alpha, Y_\beta) \in A} |Z_{\alpha, \beta}^0 - Z_{\alpha, \beta}^N|}{\sum_{(X_\alpha, Y_\beta) \in A} |Z_{\alpha, \beta}^0 - \bar{Z}_{\alpha, \beta}^0|}$$

h メッシュ幅（本章では 0.25 km を採用した）

J 繰り返し計算回数

$J_\infty$  収束に至るまでの繰り返し計算回数

M 格子点の数

M' 湖水面に存在する格子点の数

N 削減する測定点の数

$(X_\alpha, Y_\beta)$  格子点の位置

y 推定値の誤差の尺度、具体的には  $\sum_i (z_i - \hat{z}_i)^2 / \sum_i (z_i - \bar{z})^2$

$z_i$  測定点 i の測定値

$\bar{z}$   $\sum_i z_i / n$ 、ただし n は測定点の数

$\hat{z}_i$  補間によって得られる測定点 i における推定値

$\epsilon_s$  J 回の繰り返し計算の解と J + 1 回のそれとの差の平均、逐次誤差平均値と呼ぶ

$\epsilon_m$  J 回の繰り返し計算の解と J + 1 回のそれとの差の最大、逐次誤

差最大値と呼ぶ

$\eta_a$  J 回の繰り返し計算の解と最終解との差の平均、平均誤差と呼ぶ

$\eta_m$  J 回の繰り返し計算の解と最終解と差の最大、最大誤差と呼ぶ

$\omega$  SOR の加速パラメータ

$\sigma$  スプライン・パラメータ、弾性板の張力による伸びのエネルギーと歪エネルギーの係数の比

## 参 考 文 献

### 第 I 編

#### 第 1 章

- 1) 藤原正弘「環境公害データベースの作成と問題点」, 学術情報システム開発委員会(〇委員会)報告, 文部省科学研究費による特定研究「情報システムの形成過程と学術情報の組織化」(1978)
- 2) 大気汚染防止法(昭和43年6月10日、法律97), 第22条
- 3) 環境庁大気保全局「環境大気汚染常時監視保守管理マニュアル」(1979)
- 4) 東京都公害局監視部「環境監視業務等技術管理標準書」(1978)
- 5) 原田佳行「大気環境データ処理における異常値検出法」, 公害と対策, Vol.15, №2(1979)
- 6) 藤原正弘, 大島高志, 松本幸雄「環境公害データの組織化とデータ処理」, 衛生工学研究シンポジウム講演論文集, 京都大学研究シンポジウム講演論文集, 京都大学工学部衛生工学科創立20周年記念事業実行委員会(1978)
- 7) 藤原正弘, 大島高志, 松本幸雄「環境データベースの作成と利用」, 土木学会第33回年次学術講演会講演概要集第2部(1978)
- 8) 国立公害研究所環境情報部「環境データ処理システムに関する調査研究報告書」(1977)
- 9) 大島高志, 松本幸雄, 藤原正弘「国および地方自治体における大気環境データ処理システムの現状」, 公害と対策, vol 15, №2(1979)
- 10) 松本幸雄, 藤原正弘, 大島高志, 姫野純子「大気環境データ処理における標準化」, 公害と対策, Vol. 15, №3(1979)
- 11) 藤原正弘「環境データ処理利用システムの問題点と課題」, 公害と対策, Vol. 15, №2(1979)
- 12) 藤原正弘, 松本幸雄, 大島高志, 姫野純子「大気環境データの磁気テープ仕様の標準化に関する提言」, 第18回大気汚染研究全国協議会大会講演要旨集(1977)

- 13) 藤原正弘，松本幸雄，大島高志，姫野純子「大気環境データの磁気テープ仕様の標準化に関する提言（その２）」第１９回大気汚染学会講演要旨集（１９７８）
- 14) 都市計画法（昭和４３年６月１５日、法律１００），第８条第１項第１号
- 15) 藤原正弘，宇都宮陽二郎，松本幸雄，大島高志，姫野純子「環境データベースの作成 — 大気，水質環境データファイルの作成を中心として — 」，国立公害研究所発表会予稿集（１９７８）
- 16) 藤原正弘，松本幸雄，大島高志，姫野純子「国立公害研究所における環境情報の収集管理（その２）」，第１９回大気汚染学会講演要旨集（１９７９）

## 第２章

- 1) 市川 新「水質データ収集システムの現状と問題点」，公害と対策，Vol.15 №10（１９７９）
- 2) 国包章一，真柄泰基「GEMS/Water の中間総括と今後の行方」，公害と対策，Vol. 20，№13（１９８４）
- 3) 水質汚濁防止法（昭和４５年１２月２５日、法律１３８），第１５条
- 4) 環境水質保全局監修「全国公共用水域水質年鑑（１９８５年版）」，芙蓉情報センター（１９８５）
- 5) 水質汚濁防止法（昭和４５年１２月２５日、法律１３８），第１６条
- 6) 川原 浩「公共用水域における水質の自動計測」，水質汚濁研究，Vol. 18，№2（１９８５）
- 7) 建設省河川局監修「１９８３日本河川水質年鑑」，（社）日本河川協会編（１９８４）
- 8) （社）日本河川協会編「河川便覧１９８４」（１９８４）
- 9) 山村勝美「上水道の水質管理上の諸問題とその技術的対応に関する二・三の研究」，京都大学学位論文
- 10) 環境庁告示５９号「水質汚濁に係る環境基準について」，昭和４６年１２月２８日

11) 廃棄物の処理及び清掃に関する法律（昭和45年12月25日，法律137）

第2条

12) 厚生省水道環境部環境整備課監修「54年版廃棄物処理事業実態調査統計資料」，社団法人全国都市清掃会議（1980）

13) 「産業廃棄物行政組織等調査結果，1部，2部，3部」毎年度版，厚生省水道環境部産業廃棄物対策室

14) 藤原正弘「産業廃棄物対策の展望」，PPM，Vol. 13，№1（1982）

15) 横浜市環境事業局内部資料

16) 大分県，山口県，愛媛県廃棄物担当部局内部資料

17) 広域臨海環境整備センター法（昭和56年6月10日，法律76），第3条，第4条

18) 藤原正弘「広域廃棄物埋立処分場計画に伴う諸問題」，環境技術，Vol. 11，№10，703～709（1982）

19) 大阪湾広域臨海環境整備センター「昭和58年度広域処理場整備事業情報管理システム調査報告書」（1984）

20) 田中 勝，池口 孝，金子光美，河村清史「廃棄物最終処分場の事前チェックシステムの開発に関する研究」，環境庁一括計上研究，昭和57年度分（1983）

21) 神奈川県，兵庫県，京都市公害担当部局内部資料

22) 「産業廃棄物に係る情報管理システム導入手法に関する調査報告書」，厚生省水道環境部産業廃棄物対策室（1984）

23) 「産業廃棄物情報管理システム概略設計調査報告書」，厚生省水道環境部産業廃棄物対策室（1985）

### 第3章

1) 「学術情報システム開発委員会（O委員会）報告，文部省科学研究費による特定研究（1979）

2) 町田行弘，山崎 博「東京都における大気汚染情報処理システム」，公害と対策，Vol.15，№2（1979）

- 3) 原田佳行「大気環境データ処理における異常値検出法」，公害と対策，  
Vol.15, №2 (1979)
- 4) 「国立公害研究所における環境データベースの概要」，国立公害研究所  
(1978)
- 5) 「京都市公害市公害総合管理システム報告書」，京都市環境システム研究会  
(1976)
- 6) 「環境データベースのユーザガイド(第1版)」，国立公害研究所(1978)
- 7) 「環境科学に関する研究現況調査報告」，国立公害研究所(1976)
- 8) 「国際環境情報源照会システム(案内パンフレット)」，国立公害研究所  
(1976)
- 9) 「INFOTERRA 国内情報源台帳(第6版補冊)」，国立公害研究所環境情  
報部(1985)
- 10) 藤原正弘，大島高志，姫野純子「諸外国の環境情報システムについて(1)」，  
公害と対策，Vol.15, №4 (1979)
- 11) 姫野純子，大島高志，藤原正弘「諸外国の環境情報システムについて(2)」，  
公害と対策，Vol.15, №5 (1979)
- 12) 大島高志，姫野純子，藤原正弘「諸外国の環境情報システムについて(3)」，  
公害と対策，Vol.15, №6 (1979)
- 13) Federal Environmental Data :A Directory of Selected Sources,  
National Science Foundation , (1977)
- 14) Environmental Information System Directory , EPA , (1976)
- 15) SAROAD System Description and Data Coding Manual, Public  
Health Service, U.S. Department of Health, Education and Welfare  
(1968)
- 16) SAROAD User's Manual, EPA (1971)
- 17) SAROAD Parameter Coding Manual, EPA (1971)
- 18) SAROAD Station Coding Manual, EPA (1972)
- 19) SAROAD Terminal User's Manual, EPA (1973)



- 20) STORET パンフレット, EPA
- 21) STORET User Handbook, EPA
- 22) STORET Handbook, Supplementary Manual Vol.1, Data Entry,  
EPA
- 23) STORET Handbook, Supplementary Manual Vol.2, Advanced Retrieval  
EPA
- 24) WATDOC 関係文献 ( Terminal Selection Guide, Who to contact at  
WATDOC, What is WATDOC, Data Base ), Water Resources Document  
Reference Center ( 1977 )
- 25) Jabres-Bercht, Umwelt Bundesamt ( 1977 )
- 26) Environmental Information System in Sweden, Swedish Council  
of Environmental Information ( 1978 )

#### 第4章

- 1) 鈴木道夫「データベース入門」, データベースシステム, 共立出版 ( 1973 )
- 2) 藤原正弘, 大島高志, 松本幸雄「環境公害データの組織化とデータ処理」,  
衛生工学研究シンポジウム講演論文集, 京都大学工学部衛生工学科創立20周年  
記念事業実行委員会 ( 1978 )
- 3) 「環境データベースのユーザガイド ( 第2版 )」, 国立公害研究所 ( 1979 )
- 4) 藤原正弘, 宇都宮陽二郎, 松本幸雄, 大島高志, 姫野純子「環境データベース  
の作成 — 大気・水質環境データファイルの作成を中心として —」, 国立  
公害研究所研究発表会予稿集 ( 1978 )
- 5) 藤原正弘, 松本幸雄, 大島高志, 姫野純子「国立公害研究所における環境  
情報の収集・管理 ( その2 )」, 第19回大気汚染学会講演要旨集 ( 1978 )
- 6) 藤原正弘, 大島高志, 松本幸雄「環境データベースの作成を利用」, 土木学  
会第33回年次学術講演会講演概要集第2部 ( 1978 )
- 7) 松本幸雄, 藤原正弘, 大島高志, 姫野純子「大気環境データのメッシュデ  
ータ化」, 第18回大気汚染研究全国協議会大会講演要旨集 ( 1977 )

## 第Ⅱ編

### 第5章

- 1) 藤原正弘「大気汚染データの自動チェックシステムに関する研究(1)」，環境技術，Vol. 9，No. 10，19～22（1980）
- 2) 藤原正弘「大気汚染データの自動チェックシステムに関する研究(2)」，環境技術，Vol. 9，No. 11，12～22（1980）
- 3) 原田佳行「大気環境データ処理における異常値検出法」，公害と対策，Vol. 15，No. 2，35～41（1979）
- 4) 茂木正一「埼玉県における大気汚染情報処理システム」，公害と対策，Vol. 15，No. 2，20～28（1979）
- 5) 松本幸雄，藤原正弘，大島高志，姫野純子「大気環境データ処理における標準化」，公害と対策，Vol. 15，No. 3，69～80（1979）
- 6) 添田 喬，大松 繁，菊地 章「Adaptive Digital Filter による異常値検知」，徳島大学工学部報告書（1979）
- 7) 藤原正弘，姫野純子，松本幸雄，大島高志「大気環境データ処理におけるデータチェックの自動化」，京大環境衛生工学研究会第1回シンポジウム講演論文集，111～116（1979）
- 8) 姫野純子，松本幸雄，大島高志，藤原正弘「大気環境モニタリングデータの異常値の自動検出法」，公害と対策，Vol. 15，No. 8，47～51（1979）
- 9) 国立公害研究所報告書「環境公害データ処理システムに関する調査研究報告書」，国立公害研究所（1978）
- 10) 藤原正弘，大島高志，松本幸雄「環境公害データの組織化とデータ処理」，衛生工学研究シンポジウム講演論文集，京都大学工学部衛生工学科創立20周年記念事業実行委員会，32～36（1978）

### 第6章

- 1) 藤原正弘，大井 紘，新藤純子「大気汚染濃度分布パターンのスプライン法による推定法」，環境技術，Vol. 11，No. 9，643～649（1982）

- 2) 大井 紘,新藤純子,藤原正弘「大気汚染測定系の総合評価のための基礎的研究 — 地上測定局網の評価と構成」, 国立公害研究所報告 R-39-'83, 145~167 (1983)
- 3) W.R.Goodin G.J.McRae and J.H.Seineld "A Comparison of Interpolation Methods for Sparse Data : Application to Wind and Concentration Fields " Journal of Applied Meteorology Vol. 18, №6, 761~771, (1979)
- 4) 渡津秀孝,近藤 啓,富田 豊,大恵俊一郎「スプライン補間による大気汚染評価とその精度について」, 徳島大学工学部研究報告, №25, 59~71 (1980)
- 5) 篠原正則,内藤正明「スプライン法による大気汚染2次元分布の推定」, 環境情報科学, Vol.9, №4, 62~64 (1980)
- 6) 大西行雄「スプライン法を用いた2次元補間について」, Journal of the Oceanographical Society of Japan, Vol.31, №6, 259~264 (1975)

## 第7章

- 1) 藤原正弘,大井 紘,新藤純子「大気汚染測定局の最適配置 — 既存システムの検討」, 環境技術, Vol.11, №11, 788~798 (1982)
- 2) 大井 紘,新藤純子,藤原正弘「大気汚染測定系の総合評価のための基礎的研究 — 地上測定局網の評価と構成」, 国立公害研究所報告, R-39-'83, 145~167 (1983)
- 3) M.Naito and M.Ochiai "On Optimal Allocation of Air Monitoring Station", International Seminar on Air Quality Management and Related Energy Policy, Seoul Korea (1982)
- 4) 内藤正明「環境モニタリングシステムの適正化」, 国立公害研究所報告 B-10-79, 23~31, (1979)
- 5) 落合美紀子「環境モニタリング・ステーションの適正配置について」, 環境

## 第8章

- 1) 藤原正弘, 宗宮 功, 津野 洋, 藤井滋穂「水質汚濁濃度分布パターンのスプライン法による推定と合理的測定点配置の検討」, 水質汚濁研究, Vol. 8, № 2, 100～109 (1985)
- 2) 「琵琶湖南湖水質調査結果」, 京都大学工学部衛生工学教室水質工学研究室 (1981)
- 3) 「日本近海海洋汚染実態調査 — 海洋の水質測定技術向上に関する調査 —」環境庁水質保全局 (1982)
- 4) 戸川隼人「マトリクスの数値計算」, オーム社, 68 (1971)
- 5) 大西行雄「スプライン法を用いた2次元補間について」, Journal of Oceanographical Society of Japan, Vol. 31, № 6, 259～264 (1975)
- 6) 篠原正則, 内藤正明「スプライン法による大気汚染2次元分布の推定」, 環境情報科学, Vol. 9, № 4, 62～64 (1980)
- 7) 松岡 譲, 内藤正明「水質観測点の適正配置に関するシステム解析」, 国立公害研究所報告第48号 (1983)
- 8) Clasen J. (ed.) "Cooperative Programme for Monitoring of Inland Waters (Eutrophic Control), Shallow Lakes and Reservoirs, OECD Final Report, Vol. 2, (1976)